Architecting an enterprise financial management model: leveraging multi-head attention mechanism-transformer for user information transformation

Wan Yu Corresp., [1] , **Hanan Aljuaid** [2]

[1] Huanghe Science and Technology University, Zhengzhou, Henan, China

[2] Computer Sciences Department, College of Computer and Information Sciences, Princess Nourah bint Abdulrahman University (PNU), P.O. Box 84428, Riyadh 11671, Saudi Arabia

Corresponding Author: Wan Yu
Email address: amanda19881109@163.com

Financial management serves as a vital information system crucial for enterprise development. However, prevailing methods often struggle to handle the diverse information streams in financial management effectively. This paper introduces an enterprise financial management approach centered on user information signal conversion. Initially, the method enhances the Transformer network and self-attention mechanism to extract user and financial features. Subsequently, an alignment method based on reinforcement learning is proposed to reconcile the disparity between financial and user information, enhancing semantic alignment. Lastly, a signal conversion method leveraging generative adversarial networks harnesses user information for financial management, thereby elevating the efficiency of financial operations. Experimental findings substantiate the efficacy of our approach, achieving an mAP score of 81.9%. This outperforms existing methods and significantly enhances the execution performance of financial management systems.

Architecting an Enterprise Financial Management Model: Leveraging Multi-Head Attention Mechanism-Transformer for User Information Transformation

Wan Yu[1*], Hanan Aljuaid[2]

[1] Huanghe Science and Technology University, Zhengzhou 450000, Henan ,China
[2] Computer Sciences Department, College of Computer and Information Sciences, Princess Nourah bint Abdulrahman University (PNU), P.O. Box 84428, Riyadh 11671, Saudi Arabia.

Corresponding Author:
Wan Yu
Email address: amanda19881109@163.com

Abstract:

Financial management serves as a vital information system crucial for enterprise development. However, prevailing methods often struggle to handle the diverse information streams in financial management effectively. This paper introduces an enterprise financial management approach centered on user information signal conversion. Initially, the method enhances the Transformer network and self-attention mechanism to extract user and financial features. Subsequently, an alignment method based on reinforcement learning is proposed to reconcile the disparity between financial and user information, enhancing semantic alignment. Lastly, a signal conversion method leveraging generative adversarial networks harnesses user information for financial management, thereby elevating the efficiency of financial operations. Experimental findings substantiate the efficacy of our approach, achieving an mAP score of 81.9%. This outperforms existing methods and significantly enhances the execution performance of financial management systems.

Keywords: Financial management; Transformer; Reinforcement Learning; GAN

1 Introduction

The great progress of development requires enterprises to have efficient financial management capabilities to adapt to quick business transaction processing. A good financial management system can help enterprises quickly meet the market demand and quickly deploy various resources throughout the enterprise [1,2]. Therefore, the study of efficient and fast enterprise financial management systems has extremely high application value.

In addition, studying enterprise financial management helps assess and manage all kinds of risks faced by enterprises, including market risks, credit risks, liquidity risks, etc., to ensure steady enterprise development. Further analysis of financial data can provide decision support for enterprise leadership, help formulate strategic planning, optimize business models, and choose appropriate development directions [3]. Studying financial management helps enterprises make wise investment and financing decisions, choose the most suitable financing methods and investment projects for enterprise development, and protect shareholders' rights and interests to the greatest extent [4]. Studying financial management helps enterprises adapt to global competition, understand the characteristics of international markets, formulate global financial strategies, and expand international markets [5]. Therefore, the study of enterprise financial management has extremely high theoretical value.

Enterprise financial management is a complex and multi-level research in which there are some difficulties and challenges [6-8]. 1) The quality of financial data is uneven and diverse: Different enterprises may have large differences in the format, structure and specification of financial data, which leads to difficulties in data integration and analysis. 2) Finance is multi-layered: enterprise financial management involves multi-level decisions. Carrying out reasonable information transmission and decision coordination at different levels to support the realization of the enterprise's overall financial goals is a complex issue. 3) Financial risk and uncertainty: Enterprises are faced with various and complex types of risks regarding market, credit and, liquidity, etc. How to effectively quantify, identify and manage these risks and reduce uncertainty through financial management research is a challenging research direction [9-10].

Around the above difficulties, many scholars have carried out a series of studies. Chen et al. [11] posited that the swift advancement of IT has the potential to bolster organizational performance within the AIS and enhance the competitive edge of both enterprises and institutions. Kadim et al. [12] organized and revamped the accounting table, enabling an efficient display of accounting codes and product-related data for information subjects using system queries. This enhancement aims to optimize budget management for enterprises. Ren et al. [13] asserted that American colleges had implemented comprehensive management and information systems encompassing budgeting, funding, analysis and decision-making. With the development of signal processing, signal conversion algorithms have become an important way to quantify financial information. Gao et al. [14] proposed an enterprise financial information system based on cloud technology, which helps enterprises build a powerful, simple operation and strong business expansion information system at low cost through cloud computing, deep learning and other technologies. Sitinjak et al. [15] perfected the traditional financial information system by using the big data model based on the Meacher model.

However, the existing methods consider a single factor when conducting financial management, which makes it difficult to deal with the multimodal information in financial management. Besides, the existing methods cannot realize the rapid connection and modeling between users and financial information. To address this issue, we propose an enterprise financial management method based on user information signal conversion.

2 Related works

2.1 Signal Processing Techniques

Signal processing technology refers to a series of operations such as collecting, analyzing, processing, and extracting information, improving characteristics, recognizing patterns, and predicting future behavior of signals. These signals can be sounds, images, texts, videos, or even complex biological signals, among others. The signal processing technology is widely used in communication, audio processing, image processing, biomedicals, finance, automatic control and other fields.

In the field of signal processing, the dLMS algorithm [16, 17] usually has a slow convergence speed because the covariance matrix of the signal usually has a large eigenvalue spread. At the same time, dLMS proves that dLMS based on Newton's method (dLMSN) has a faster convergence speed than the traditional dLMS algorithm, but its cost is the need to calculate the inverse matrix, which has high algorithm complexity. Furthermore, Hua et al. [18] proposed the diffusion Preconditioned LMS (dPLMS) based on preconditioning, which uses preconditioning operations to boost the convergence rate of the algorithm and approximately achieves the stable status of the dLMSN.

Financial management is signal processing research with careful consideration of multitasks. In the context of multitask parameter estimation, it is common to partition the nodes within the network into several clusters, with each cluster consisting of nodes assigned to perform a specific parameter estimation task. Chen et al. [19] realized the collaborative estimation of multiple parameters by predicting the clustering information and the relationship between different tasks and fully considering the collaboration between clusters. Nassif et al. [20] assume that the parameters estimated by nodes in different clusters are linearly correlated, and the linear constraint relationship between the parameters to be estimated is known in advance. Then, when the step size of the dLMS algorithm is small enough, each node can collaboratively achieve the optimal estimation of any accuracy. On the contrary, Chen et al. [21] proposed that when there is no prior knowledge about different tasks in multitask scenarios, the unsupervised clustering strategy can adaptively adjust the combination strategy of each node according to the similarity of the estimation tasks of each node, so that each node can realize adaptive collaborative estimation.

2.2 Current research status of financial management

Financial management technology refers to a series of methods and tools that apply information technology and related software tools to improve, optimize and automate the financial activities of enterprises. These technologies can help enterprises more efficiently carry out financial planning, financial analysis, budget control, cost management, risk assessment, investment decisions and other aspects of the work.

Berdiev [22] believes that by using the information technology platform of the Internet, enterprises can carry out various commercial trade activities, make the form of trade more efficient, reduce the cost of commercial input, expand the scope of commercial activities, and continuously realize the diversification of commercial operations. Polzer et al. [23] held the view that the swift evolution of the Internet has transformed work dynamics. Financial management is also undergoing continuous reform with the improvement of information technology. Yermack et al. [24] believed that in the "Internet +" market environment, the financial management structure of enterprises is also under continuous reform.

Gigli et al. [25] proposed that in the process of transformation from a cash basis to an accrual basis, excessive bureaucratization of management, the complexity of organizational structure, limitations of IT system and regulatory loopholes will affect the reform and innovation of financial administration. Therefore, it is necessary to consider the establishment of a framework system from the aspects of system, organization and technology. Hoerlsberger et al. [26] put forward that the integration of IT has a great influence on universities, governments, enterprises, and management innovation not only through the construction of information but also through structural changes. To adapt to the development of digital trends, optimize and innovate in organizational structure, personnel setup and management mechanism, and establish a new management system [27]. Paterson et al. [28] proposed an accounting system based on information technology, which simplifies daily tasks and work, simplifies financial reporting and disclosure, and changes the development of the whole financial management. Cooper et al. [29] proposed an automatic input, processing and output of information data based on intelligent robots to simplify repetitive work, liberate labor, and improve work efficiency.

3. Enterprise financial management method based on user information signal conversion

The existing financial management methods are difficult to deal with the multimodal information in the system and cannot realize the rapid connection and modeling between users and financial information. In this paper, an enterprise financial management method based on user information signal conversion is proposed. Aiming to align user information and financial information, this method proposes a multimodal feature extraction method based on an improved transformer, a feature alignment method by reinforcement learning, and a signal conversion method by the generative adversarial network. Combined with the signal conversion of user information, the mutual relationship between users and finance is modeled.

3.1 Multimodal Feature extraction Method based on improved Transformer

In this paper, the Transformer-based multimodal feature extraction method is used as a feature quantification method for financial information and user information, and its codec structure has a strong receptive field, as shown in Figure 1. Transformer is a kind of codec. By calculating the similarity between tokens, it extracts semantic information and realizes the feature language expression. The Transformer network improves the malleability of LSTM, which can effectively model the relationship between two elements with a large distance and take into account the relationship mining between short-term elements. In addition, the basic unit of the Transformer is matrix operation, which avoids the cyclic calculation of elements and saves computing resources during training and testing.

Figure 1 Structure of our improved Transformer

By improving the network structure and increasing the direct information exchange between the encoder and decoder, this paper further Narrows the semantic gap caused by cross-modal conversion. In the improved Transformer network, the most critical component is the Multi-head Attention mechanism (MHSA). The multi-head attention structure ccontains multiple Self-attention. In principle, the self-attention mechanism uses the Query feature (Q) to weight the Key-Value feature group. This process involves computing a similarity matrix between the Query feature and the Key feature (K), and then using this similarity matrix to weight the Value feature (V). This is computed as follows:

$$Q = \mathrm{x}W^{Q} \tag{1}$$

$$K = \mathrm{x}W^{K} \tag{2}$$

$$V = \mathrm{x}W^{V} \tag{3}$$

$$Attention = softmax\left(\frac{QK^{T}}{\sqrt{d_k}}\right)V \tag{4}$$

where W^Q, W^K and W^V are the learnable parameters, x is the sequence feature vector, and d_k is the dimension of Q feature and K feature. The above equation is to compute the Self-attention mechanism. Furthermore, in order to effectively leverage the sequencing of financial information and augment the capabilities of the Self-attention mechanism, we introduce an enhanced self-attention mechanism, depicted in Figure 2. Subsequently, the enhanced Self-attention mechanism is employed to construct the MHSA, which is illustrated in the following equation:

$$MultiHead(Q,K,V) = \left[head_1,\ldots,head_n\right]W \tag{5}$$

$$head_i = Attention\left(Q_i,K_i,V_i\right) \tag{6}$$

Figure 2 Improved Self-attention

3.2 Feature alignment method based on reinforcement learning

Through the Transformer-based multimodal feature extraction method, we can obtain the financial features of the enterprise and the user features of the enterprise. However, these two features belong to different semantic fields, which directly leads to the subsequent feature matching is not differentiable. Moreover, there is a large gap between the evaluation index of our method and the commonly used cross-entropy loss, which cannot directly improve the model through the training of the model, resulting in the results of training and testing being difficult to be consistent, and errors will continue to accumulate in the process of alignment.

Reinforcement learning (RL) is a AI method of that focuses on how an intelligent agent can learn to make decisions in specific scenarios through interaction with the environment, to maximize the overall reward (or minimize the penalty). In reinforcement learning, an intelligent agent performs a sequence of actions and subsequently ADAPTS its approach based on the response (reward or punishment) of the environment to improve the outcome over a longer period. This learning process usually involves a trial-and-error process, where the agent gradually improves its behavior by constantly trying various strategies.

Figure 3 Signal conversion method based on reinforcement learning

The core of the feature alignment by reinforcement learning is to regard signal conversion as a reinforcement process, as shown in Figure 3, which causes the training purpose of the model to be changed to minimize the negative value of the expected reward. The formula is as follows:

$$L(\theta) = -\mathrm{E}_{w \sim p_\theta}\left[r\left(w^s\right)\right] \tag{7}$$

Since the evaluation indexes of signal conversion are not differentiable, the constrained gradient of $L(\theta)$ can be obtained based on Monte-Carlo theory, as shown in the following formula:

$$\nabla_\theta L(\theta) = -\mathrm{E}_{w \sim p_\theta}\left[r\left(w^s\right)\nabla_\theta log p_\theta\left(w^s\right)\right] \tag{8}$$

$$\nabla_\theta L(\theta) = -\mathrm{E}_{w \sim p_\theta}\left[\left(r\left(w^s\right) - b\right)\nabla_\theta log p_\theta\left(w^s\right)\right] \tag{9}$$

In the training of financial information and personal information signal transformation, b is usually defined as the reward value predicted by the model at the current time, denoted as r(w '), and its formula is as follows:

$$\frac{\partial L(\theta)}{\partial \mathrm{s_t}} = \left(r\left(w^s\right) - r\left(w^{'}\right)\right)\left(p_\theta\left(w^s|h_t\right) - 1\right) \tag{9}$$

Among them, the reward value is usually calculated using evaluation metrics to better maintain the consistency of model training and testing performance.

3.3 Signal Conversion Method by GAN

To model the relationship between users and finance, we propose a signal conversion method by generative adversarial networks.

By learning the probability distribution of financial signal data, GAN can convert the signal input by the user into a financial signal, so that it can be identified as a real financial signal through the discrimination network. The framework of the generated network is shown in Figure 4, and a CNN based on the Unet++ structure is emplyed. The structure is divided into three parts: encoding path, decoding path and skip connection. In the figure, the number in each rectangular box represents the length of the data multiplied by the number of channels. The encoding part down-samples the signal, and each step contains a convolution module and a Max pooling layer to realize the extraction of signal features. The decoding part is opposite to the encoding part, with the difference that the Max pooling layer is replaced by an upsampling layer. The skip connection part is also different. It is different from the Unet structure that directly connects the feature map of the encoding path and the decoding path, but integrates the convolution module into the skip connection, and fuses the features of the next stage of convolution to optimize the feature fusion step. The channels of the feature needs to be adjusted by the convolution, each convolution includes two one-dimensional convolution layers, the Batch normalization (BN) layer, and the ReLU activation layer. Finally, the four upsampled feature maps were fed into a one-dimensional convolutional layer to output data with the same dimension.

Figure 4 Structure of the generated network

The discriminant network is used to estimate the probability that the user data converted by the generation network is consistent with the original financial data, which is essentially a binary classifier. The network is presented in Figure 5. The number in each rectangular box in the figure represents the length × the number of channels of the data after the convolution module. The original financial signal and the generated user signal are concatenated and input into the discriminative network. The network first uses a one-dimensional convolution layer with convolution kernel size 3 and step size 1 to extract shallow features, and then uses four identical modules. In each module, it includes a 1-D convolution layer with convolution kernel size 4 and step size 2, a BN and a activation layer. Finally, a one-dimensional convolutional layer is used to convert the number of channels to 1, and the Sigmoid function is utilized to achieve the classification result. The LeakyReLU activation function is used to prevent overfitting, and the Negative slope is set to 0.2 to ensure that the gradient transfer is simpler.

Figure 5 Discriminative network structure

4 Experiment and analysis

4.1 Dataset and implement details

We employ the Enterprise Finance Dataset to test the Enterprise financial management method based on user information signal conversion. The dataset was sourced from enclosures within company financial reports submitted to the commission. The dataset is available at Enterprise Financial Dataset (zenodo.org). It encompasses financial statements, which are more condensed compared to the comprehensive financial statements and Notes dataset. The latter includes both numerical and narrative disclosures for all financial statements and accompanying notes. The information provided remains consistent with the financial reports "filed" by each registrant. The data is structured in a straightforward format, aiding users in analyzing and comparing company disclosures across time and among registrants. Additionally, the dataset incorporates supplementary fields like standardized industry classifications for companies, streamlining data utilization.

Since Transformers, reinforcement learning, and Gans are the most popular training models for big data, we will use CPU: Xeon(R) E5-2640 v4 and GPU: 4*Nvidia Tesla V100 to build the environment and train the model. The deep learning framework is Tensorflow. The experimental parameters are presented in Table 1.

Table 1 Implementation parameters

Since the enterprise financial management method is a multimodal task, we adopt the mean square Error (mAP) and F-measure as the evaluation criteria of the method, which are calculated as follows:

$$V_P = \frac{gt \cap pr}{pr} \tag{13}$$

$$V_R = \frac{gt \cap pr}{gt} \tag{14}$$

$$F = \frac{2 \times V_P \times V_R}{V_P + V_R} \tag{15}$$

$$\mathrm{mAP} = \frac{1}{N} \times \sum V_P \times V_R \tag{16}$$

where pr refers to the result of the method and gt denotes the true value present in the dataset. In addition, we evaluate the performance of the enterprise financial management by the amount of calculation, the number of model parameters, and the operation time.

Figure 6 Compare our method with others

4.2 Compare our detection method with others

First, we conduct experiments of the improved Transformer-based multimodal feature extraction method on the Enterprise Finance dataset. We select some excellent feature models, such as Transformer [30], Bert [31], Oscar [32] and VinVL [33] and DFT [34], and compare the performance. The results are presented in Figure 6 and Table 2. We can conclude that our method obtains the highest value in all evaluation metrics, which is 0.942 for recall, 0.915 for precision, 0.936 for F-measure, and 0.924 for mAP, while comparing with other algorithms. Compared with the Transformer, our method improves the mAP value of the model by more than 7%, mainly because we improve the Transformer and optimize the self-attention mechanism. Compared with BERT, our method has a lead of more than 6%. BERT is almost the same as a Transformer in principle, so the performance is comparable between the two. Oscar and VinVL are models that are pre-trained with big data and have better adaptability to multimodal tasks, and our method still obtains more than 3% improvement in the mAP score. Compared with the latest DFT method, which can obtain more than 90% of the mAP value by excellent model performance, our method still obtains about 2% advantage. For the extraction of enterprise financial features and user features, the multimodal feature extraction by the improved Transformer proposed in this paper has great advantages. Through the improved self-attention mechanism and recurrent Transformer structure, the financial features and user features can be effectively extracted and aligned.

Table 2 Compare our detection method with other methods

Figure 7 The performance of our signal process method

Then, we implement the performance test of the reinforcement learning-based feature alignment method on the dataset. Feature alignment is mainly evaluated by the alignment yield. Therefore, the conversion rate of feature alignment was used as the index in this experiment. We still compare the our method to Transformer [30], Bert [31], Oscar [32] and VinVL [33] and DFT [34]. From Figure 7, it can be found that the feature alignment method by RL achieves the highest feature alignment yield, that is, 78.6%, which reaches the most advanced level in the world. In addition, the training process is recorded, and the convergence graph is generated. In Figure 7, the training process of the proposed method is very smooth, and the lowest loss value can be obtained, which fully demonstrates the stability and scalability of the proposed method to data.

Table 3 Compare our detection method with other methods

Figure 8 The results of our method

After verifying the multimodal feature extraction method based on the improved Transformer and the signal conversion method based on reinforcement learning, we will verify our proposed signal conversion method by the GAN in the dataset. Our method is compared with some excellent models, such as ActFormer [35], Git [36], CP-GAN [37], and NF-ResNet [38]. The evaluation metrics are still Recall, Precision, F measure and mAP, and the results are in Figure 8 and Table 3. Our method achieved the highest value of 0.823 in Recall, 0.837 in Precision, and 0.826 in F measure among all evaluation metrics. Compared with ActFormer, our method improves the mAP score by more than 5%, and improves the F measure score by 5%. Compared with Git, our method has more than 3% F-measure lead and 3.4% mAP value boost. Compared with CP-GAN, our method leads in all aspects, and all evaluation indexes are higher than 2%. Finally, compared with NF-ResNet, our method improves the mAP score by 1.4% and the F measure score by 1.2%. Since the GAN model is based on Unet and contains fewer layers and feature processing steps, the signal transformation method by the GAN proposed in this paper has far superior performance to other methods.

Finally, we integrate the above three methods to implement an end-to-end enterprise financial management system, and take the number of model parameters, inference time, Flops and training time as system evaluation indicators. As shown in Figure 9, the proposed method can obtain the least training and testing time and the least number of model parameters.

Figure 9 Model efficiency comparison with other methods

5 Conclusion

In response to the requirements of advancing enterprise information systems, this paper suggests an approach to enterprise financial management. This method centers around converting user information signals, effectively addressing the challenge posed by managing diverse modalities of information within the financial domain.By proposing a multimodal feature extraction by an improved Transformer, a feature alignment method by reinforcement learning, and a signal conversion method by a generative adversarial network, feature extraction, feature alignment, and signal conversion are improved in turn. These enhancements serve to bridge the semantic gap between financial information and user data, enabling a comprehensive modeling of their interrelation. Ultimately, this paves the way for optimizing the financial management system. Experiments prove that our method can obtain a mAP score of 81.9%, which can improve the performance of enterprise financial management systems.

Reference

[1] Sukenti S. Financial Management Concepts: A Review[J]. Journal of Contemporary Administration and Management (ADMAN), 2023, 1(1): 13-16.

[2] Sitinjak C, Johanna A, Avinash B, et al. Financial Management: A System of Relations for Optimizing Enterprise Finances–a Review[J]. Journal Markcount Finance, 2023, 1(3): 160-170.

[3] Winarno A, Agustina Y, Vinola R. Developing website-based school financial administrative management system during covid-19 pandemic[J]. International Journal of Business, Economics and Law, 2020, 22(1): 167-172.

[4] Bustani B, Khaddafi M, Ilham R N. Regional Financial Management System of Regency/City Regional Original Income In Aceh Province Period Year 2016-2020[J]. International Journal of Educational Review, Law And Social Sciences (IJERLAS), 2022, 2(3): 459-468.

[5] Klapper L, Lusardi A. Financial literacy and financial resilience: Evidence from around the world[J]. Financial Management, 2020, 49(3): 589-614.

[6] Mazur N, Khrystenko L, Pasztorova J, et al. Improvement of controlling in the financial management of enterprises[J]. TEM Journal-Technology, Education, Management, Informatics, 2021.

[7] Cui R. Iot financial management system for energy enterprise management risk and prevention and control strategy under the background of double carbon[J]. 3c Empresa: investigacion y pensamiento critico, 2023, 12(2): 144-160.

[8] Krilova V V. Financial Management of an Industrial Enterprise in the Context of Digital Transformation[J]. Юго-Западного государственногоуниверситета, 2023:166.

[9] Abad-Segura E, Gonzalez-Zamar M D, Lopez-Meneses E, et al. Financial technology: review of trends, approaches and management[J]. Mathematics, 2020, 8(6): 951.

[10] Aifuwa H O, Embele K. Board characteristics and financial reporting[J]. Journal of Accounting and Financial Management, 2019, 5(1): 30-44.

[11] Chen X, Metawa N. Enterprise financial management information system based on cloud computing in big data environment[J]. Journal of Intelligent & Fuzzy Systems, 2020, 39(4): 5223-5232.

[12] Kadim A, Sunardi N. Financial management system (QRIS) based on UTAUT model approach in jabodetabek[J]. International Journal of Artificial Intelligence Research, 2023, 6(1.1).

[13] Ren S. Optimization of enterprise financial management and decision-making systems based on big data[J]. Journal of Mathematics, 2022, 2022: 1-11.

[14] Gao J. Analysis of enterprise financial accounting information management from the perspective of big data[J]. International Journal of Science and Research (IJSR), 2022, 11(5): 1272-1276.
[15] Sitinjak C, Johanna A, Avinash B, et al. Financial Management: A System of Relations for Optimizing Enterprise Finances–a Review[J]. Journal Markcount Finance, 2023, 1(3): 160-170.
[16] Nassif R, Richard C, Chen J, et al. A graph diffusion LMS strategy for adaptive graph signal processing[C]//2017 51st Asilomar Conference on Signals, Systems, and Computers. IEEE, 2017: 1973-1976.
[17] Nassif R, Richard C, Chen J, et al. Distributed diffusion adaptation over graph signals[C]//2018 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP). IEEE, 2018: 4129-4133.
[18] Hua F, Nassif R, Richard C, et al. A preconditioned graph diffusion LMS for adaptive graph signal processing[C]//2018 26th European Signal Processing Conference (EUSIPCO). IEEE, 2018: 111-115.
[19] Chen J, Richard C, Sayed A H. Multitask diffusion adaptation over networks[J]. IEEE Transactions on Signal Processing, 2014, 62(16): 4129-4144.
[20] Nassif R, Richard C, Ferrari A, et al. Distributed learning over multitask networks with linearly related tasks[C]//2016 50th Asilomar Conference on Signals, Systems and Computers. IEEE, 2016: 1390-1394.
[21] Chen J, Richard C, Sayed A H. Diffusion LMS over multitask networks[J]. IEEE Transactions on Signal Processing, 2015, 63(11): 2733-2748.
[22] Berdiev A O. Improvement of Financial Management Mechanisms of Investment Activity[J]. EUROPEAN JOURNAL OF BUSINESS STARTUPS AND OPEN SOCIETY, 2023, 3(5): 62-68.
[23] Polzer T, Nolte I M, Seiwald J. Gender budgeting in public financial management: a literature review and research agenda[J]. International Review of Administrative Sciences, 2023, 89(2): 450-466.
[24] Yermack D. Donor governance and financial management in prominent US art museums[J]. Journal of Cultural Economics, 2017, 41(3): 215-235.
[25] Gigli S, Mariani L, Trivellato B. Assimilating accounting innovations: lessons from Italian public universities[J]. Journal of Accounting & Organizational Change, 2018, 14(4): 381-401.
[26] Hoerlsberger M. Innovation management in a digital world[J]. Journal of Manufacturing Technology Management, 2019, 30(8): 1117-1126.
[27] Mohammad A A. An exploration of accounting information system's role in SMEs failure triangular[J]. International Journal of Agile Systems and Management, 2018, 11(2): 155-178.
[28] Paterson A S, Wilson M J. The Routledge companion to accounting information systems, edited by Martin Quinn and Erik Strauss, published in 2018 by Routledge Taylor & Francis Group: London and New York[C]//Accounting Forum. Taylor & Francis, 2018, 42(3): 277-280.
[29] Cooper L A, Holderness Jr D K, Sorensen T L. Robotic process automation in public accounting[J]. Accounting Horizons, 2019, 33(4): 15-35.

[30] Vaswani A, Shazeer N, Parmar N. Attention is all you need[J]. Advances in neural information processing systems, 2017, 30.
[31] Deepa M D. Bidirectional encoder representations from transformers (BERT) language model for sentiment analysis task[J]. Turkish Journal of Computer and Mathematics Education (TURCOMAT), 2021, 12(7): 1708-1721.
[32] Li X, Yin X, Li C. Oscar: Object-semantics aligned pre-training for vision-language tasks[C]//Computer Vision–ECCV 2020: 16th European Conference, Glasgow, UK, August 23–28, 2020, Proceedings, Part XXX 16. Springer International Publishing, 202:121-137.
[33] Zhang P, Li X, Hu X. Vinvl: Revisiting visual representations in vision-language models[C]//Proceedings of the IEEE/CVF conference on computer vision and pattern recognition. 2021: 5579-5588.
[34] Zhang J, Xie Y, Ding W. Cross on cross attention: Deep fusion transformer for image captioning[J]. IEEE Transactions on Circuits and Systems for Video Technology, 2023.
[35] Xu L, Song Z, Wang D. ActFormer: A GAN-based Transformer towards General Action-Conditioned 3D Human Motion Generation[C]//Proceedings of the IEEE/CVF International Conference on Computer Vision. 2023: 2228-2238.
[36] Wang J, Yang Z, Hu X. Git: A generative image-to-text transformer for vision and language[J]. arXiv preprint arXiv:2205.14100, 2022.
[37] Skariah D, Thomas J. Review of Speech Enhancement Methods using Generative Adversarial Networks[C]//2023 International Conference on Control, Communication and Computing (ICCC). IEEE, 2023: 1-4.
[38] Khan, F. A., Jamjoom, M., Ahmad, A., & Asif, M. (2022). An analytic study of architecture, security, privacy, query processing, and performance evaluation of database-as-a-service. *Transactions on emerging telecommunications technologies*, *33*(2), e3814.

Figure 1

Structure of our improved Transformer

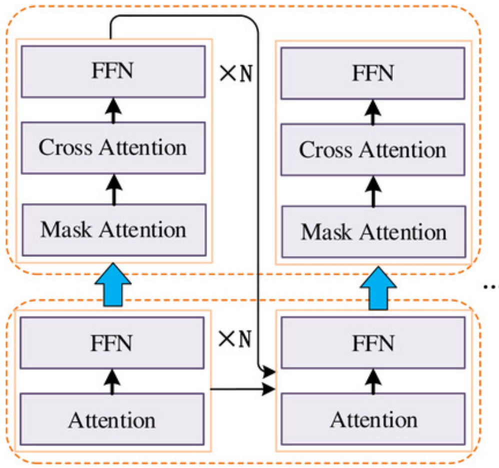
FFN
Cross Attention
Mask Attention
×N
FFN
Cross Attention
Mask Attention
...
FFN
Attention
×N
FFN
Attention

Figure 2

Improved Self-attention

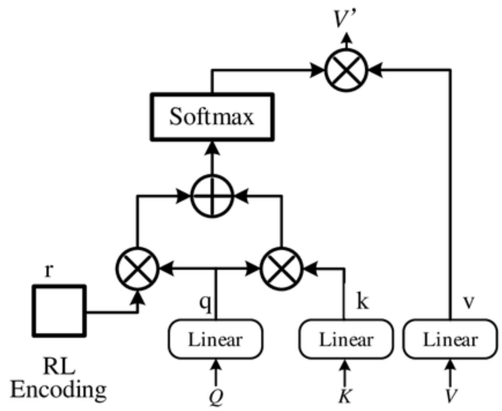
V'
Softmax
r
q
k
v
RL
Encoding
Linear
Linear
Linear
Q
K
V

Figure 3

Signal conversion method based on reinforcement learning

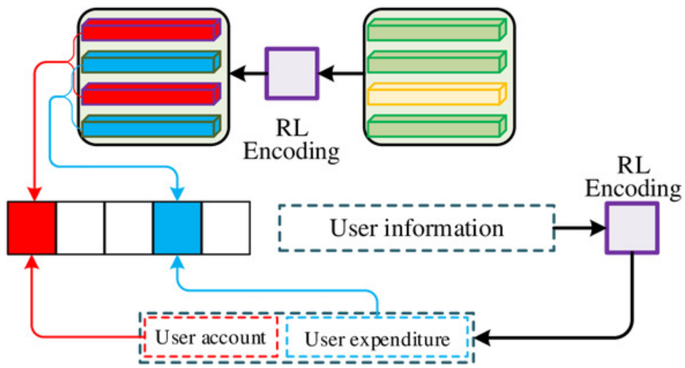
RL
Encoding
RL
Encoding
User information
User account
User expenditure

Figure 4

Structure of the generated network

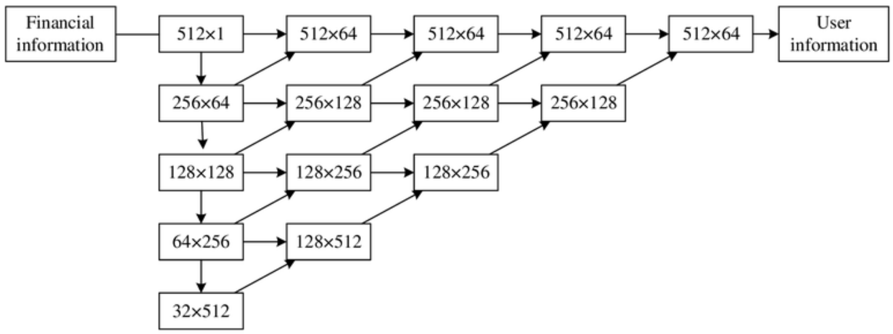
Financial information
512×1
512×64
512×64
512×64
512×64
User information
256×64
256×128
256×128
256×128
128×128
128×256
128×256
64×256
128×512
32×512

Figure 5

Discriminative network structure

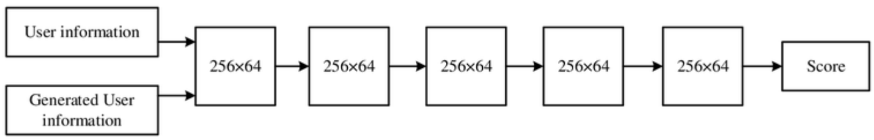
User information
Generated User information
256×64
256×64
256×64
256×64
256×64
Score

Figure 6

Compare our method with others

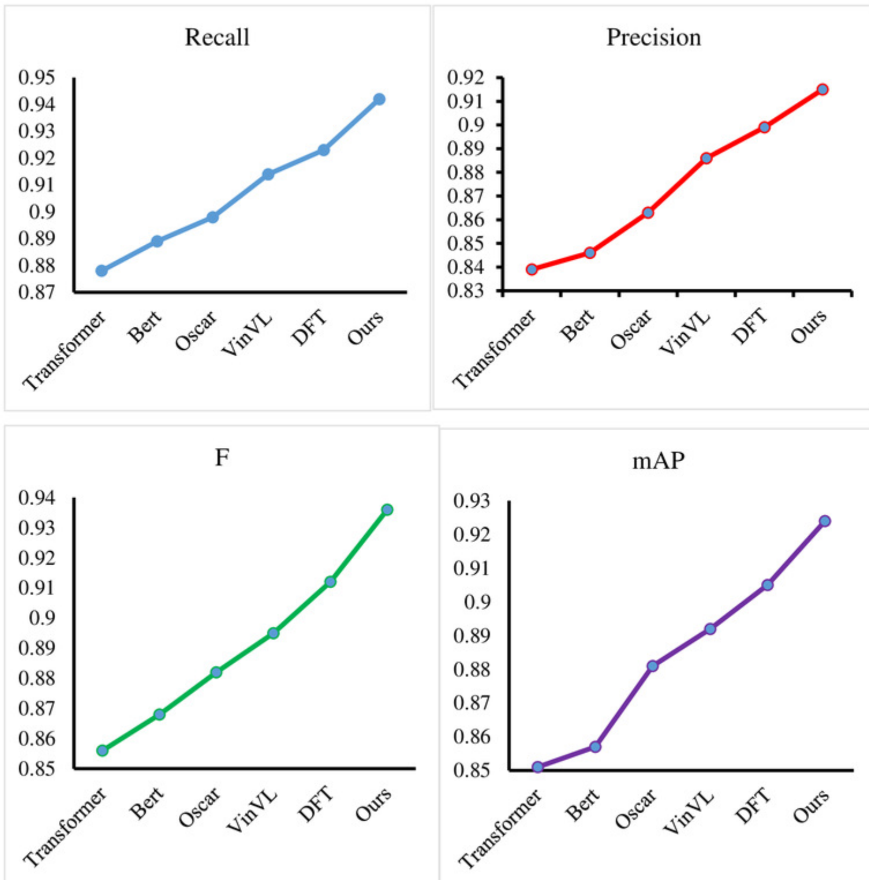
Recall
0.95
0.94
0.93
0.92
0.91
0.9
0.89
0.88
0.87
Transformer
Bert
Oscar
VinVL
DFT
Ours
Precision
0.92
0.91
0.9
0.89
0.88
0.87
0.86
0.85
0.84
0.83
Transformer
Bert
Oscar
VinVL
DFT
Ours
F
0.94
0.93
0.92
0.91
0.9
0.89
0.88
0.87
0.86
0.85
Transformer
Bert
Oscar
VinVL
DFT
Ours
mAP
0.93
0.92
0.91
0.9
0.89
0.88
0.87
0.86
0.85
Transformer
Bert
Oscar
VinVL
DFT
Ours

Figure 7

The performance of our signal process method

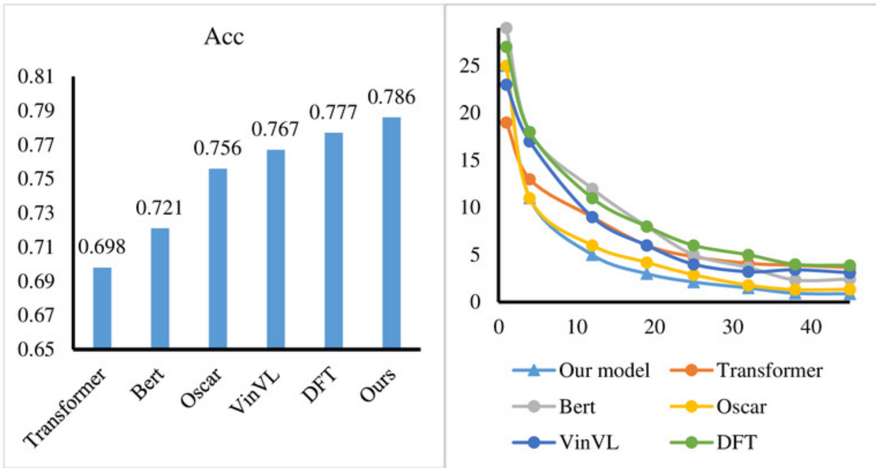
Acc
0.81
0.79
0.77
0.75
0.73
0.71
0.69
0.67
0.65
0.698
0.721
0.756
0.767
0.777
0.786
Transformer
Bert
Oscar
VinVL
DFT
Ours
25
20
15
10
5
0
0
10
20
30
40
Our model
Transformer
Bert
Oscar
VinVL
DFT

Figure 8

The results of our method

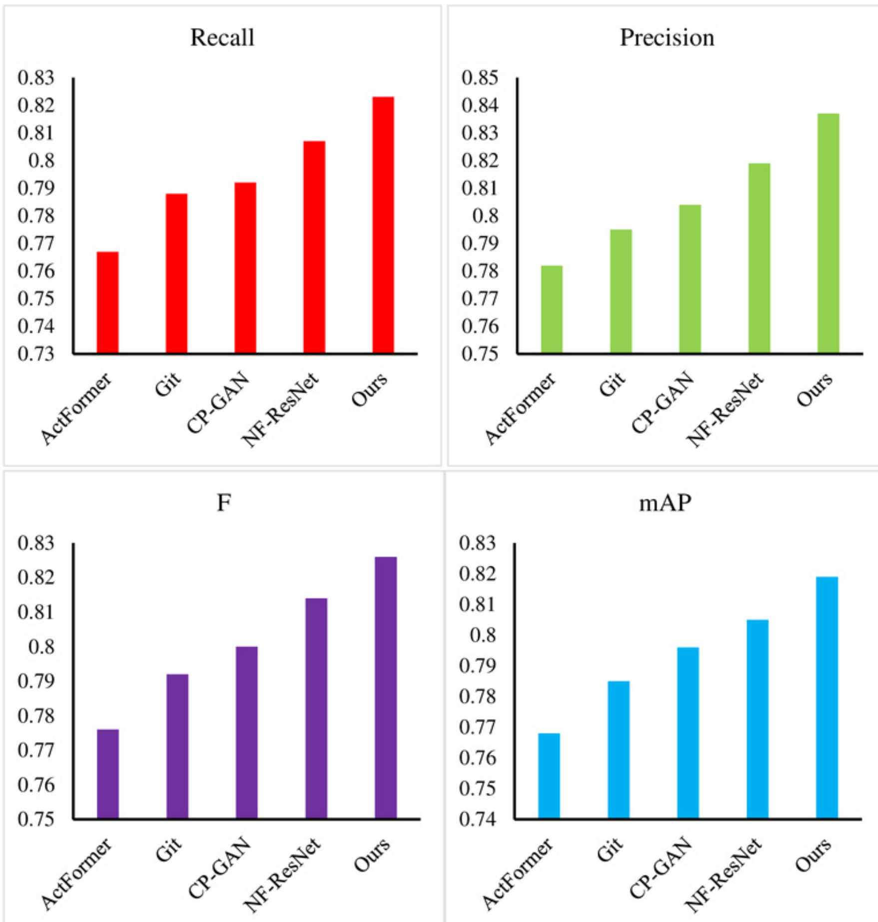
Recall
0.83
0.82
0.81
0.8
0.79
0.78
0.77
0.76
0.75
0.74
0.73
ActFormer
Git
CP-GAN
NF-ResNet
Ours
Precision
0.85
0.84
0.83
0.82
0.81
0.8
0.79
0.78
0.77
0.76
0.75
ActFormer
Git
CP-GAN
NF-ResNet
Ours
F
0.83
0.82
0.81
0.8
0.79
0.78
0.77
0.76
0.75
ActFormer
Git
CP-GAN
NF-ResNet
Ours
mAP
0.83
0.82
0.81
0.8
0.79
0.78
0.77
0.76
0.75
0.74
ActFormer
Git
CP-GAN
NF-ResNet
Ours

Figure 9

Model efficiency comparison with other methods

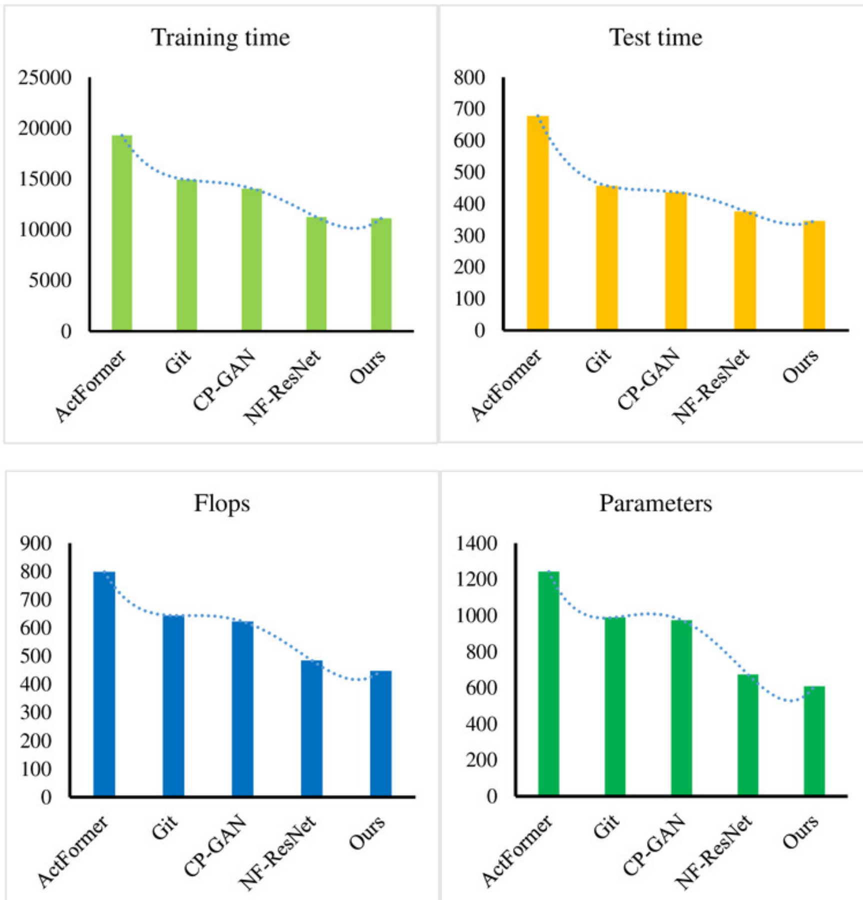
Training time
25000
20000
15000
10000
5000
0
ActFormer
Git
CP-GAN
NF-ResNet
Ours
Test time
800
700
600
500
400
300
200
100
0
ActFormer
Git
CP-GAN
NF-ResNet
Ours
Flops
900
800
700
600
500
400
300
200
100
0
ActFormer
Git
CP-GAN
NF-ResNet
Ours
Parameters
1400
1200
1000
800
600
400
200
0
ActFormer
Git
CP-GAN
NF-ResNet
Ours

Table 1(on next page)

Implementation parameters

Parameters	value
Initial learning rate	8×10^{-4}
Epoch	80
Batch-size	40
Decay	0.95
Gradient descent method	SGD
Image input size	380 x 380
Image feature dimension	1024

Table 2(on next page)

Compare our detection method with other methods

Methods	Recall	Precision	F	mAP
Transformer	0.878	0.839	0.856	0.851
Bert	0.889	0.846	0.868	0.857
Oscar	0.898	0.863	0.882	0.881
VinVL	0.914	0.886	0.895	0.892
DFT	0.923	0.899	0.912	0.905
Ours	0.942	0.915	0.936	0.924

Table 3 (on next page)

Compare our detection method with other methods

Methods	Recall	Precision	F	mAP
ActFormer	0.767	0.782	0.776	0.768
Git	0.788	0.795	0.792	0.785
CP-GAN	0.792	0.804	0.800	0.796
NF-ResNet	0.807	0.819	0.814	0.805
Ours	0.823	0.837	0.826	0.819