SUGISAKA MASANORI

Journal of Robotics, Networking and Artificial Life
Vol. 11(1); May (2025), pp. 83–88
ISSN (Online): 2352-6386; ISSN (Print): 2405-9021
https://alife-robotics.org/jrnal.html

Research Article

Recognition of Guqin Music Notation of Jianzi Pu by Deep Learning Methods

Takashi Kuremoto[1], Kazuma Fujino[1], Hirokazu Takahashi[1], Shun Kuremoto[2, 3], Mamiko Koshiba[2], Hiroo Hieda[3], Shingo Mabu[2]

[1]*Nippon Institute of Technology, Gakuendai 4-1, Miyashiro-Machi, Saitama, 345-8501, JAPAN*
[2]*Yamaguchi University, Tokiwadai 2-16-1, Ube, Yamaguchi, 755-8611, JAPAN*
[3]*Institute of Future Engineering, Fukagawa 2-6-11, Koto-ku, Tokyo, 135-8473 JAPAN*

ARTICLE INFO

Article History

Received 21 November 2023
Accepted 21 April 2025

Keywords

Chinese seven-string zither
Deep learning
Guqin notation
Sen-O-So
VGG16
YOLOv5

ABSTRACT

The music notation of Guqin (“古琴”, Chinese seven-string zither) named “Jianzi Pu (減字譜, simplified music notation of Guqin)” was invented at the Middle of A.D. 700, and Guqin music remained more than 600, however, only about 100 of them are played in nowadays. The reason is that the handwritten “Jianzi Pu” is hard to be understood even for experts or professional Guqin players. In this study, we applied deep learning methods such as VGG16 and YOLOv5 to the recognition of a Guqin notation “Sen-O-So” (仙翁操, Melody of the Immortal Elder). Firstly, we created a dataset including 55 kinds of single characters of Sen-O-So in 4,951 images from 23 versions found on the Internet and obtained by data augmentation, i.e., image processing such as rotation, enlargement (zoom-in), reduce (zoom-out), various filtering, etc. Secondly, we compared the recognition rates of VGG16 and YOLOv5 in the experiment. The average accuracies of 55 classes images by VGG16 and YOLOv5 were 87.50% and 88.47% respectively for the test data. Additionally, we created a dataset of Sen-O-So video clips to match the recognition results of single characters by YOLOv5 and realized an online ancient music restoration system development.

1. Introduction

Over 3,000 years ago in China, a seven-stringed zither known as the Guqin (古琴, Chinese seven-string zither) emerged, one of the oldest stringed instruments in the world. The Guqin is crafted with a paulownia wood front and catalpa wood back, measuring approximately 1.3 meters in length, 20 centimeters in width, and 5 centimeters in thickness, with seven strings stretched across its surface (Figs. 1 and 2). More than just a musical instrument, the Guqin has been revered as a cultural symbol in East Asian history, symbolizing character cultivation and being integral in important ceremonies. However, much of the instrument's music and literature were lost during the medieval period. Today, although 3,360 qin pieces, 130 qin notation collections, and 300 qin songs still exist in written form, only a few dozen is actively performed [1]. This limited repertoire is due to the nature of Guqin tablature, known as Jianzi Pu (減字譜, simplified music notation of Guqin) (see Figs. 3 and 4), which records pitch but not rhythm. Deciphering and interpreting this tablature for performance, a process known as “Dapu”(打譜, notation-music translation), is essential to make the ancient scores understandable within modern music theory (such as the staff notation). Specialists capable of Dapu are few worldwide (fewer than several dozen), while the number of young Guqin players in China is expanding into the hundreds of thousands, leaving much of the existing Guqin repertoire still dormant [1] [2].

The academic question driving this research is how deep learning techniques can be applied to automatically

Corresponding author E-mail: kuremoto.takashi@nit.ac.jp

interpret Jianzipu, the handwritten Guqin tablature, with the aim of reviving a vast number of currently unperformed Guqin pieces. These Guqin music are translated from Jianzi Pu to modern staff notation by experts of Guqin players, and this translation process is called “Dapu” [3]. Jianzi Pu recognition by computer has been studied since last century in China [4] [5] [6], however, the practical results are not obtained.

In our previous work [7] [8], machine learning methods such as support vector machine (SVM), deep learning models VGG16, ResNet50, and the hybrid models VGG16 with SVM, ResNet50 with SVM were adopted to be classifiers of single character of Jianzi Pu. In the comparison experiments using a dataset of a well-known Guqin music “Sen-O-So” (仙翁操, Melody of the Immortal Elder), VGG16 with SVM had the highest average accuracies to 15 kinds of single characters data which were the first and the second line of Sen-O-So Jianzi Pu in 1,500 images, i.e., 99.11% in training and 88.33% for unknown data.

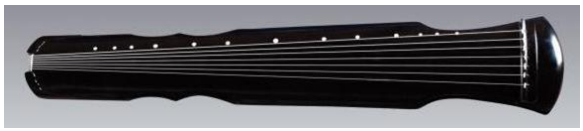


Fig. 1 A musical instrument: Guqin (7 strings Qin).

Fig. 2 Playing Guqin (Player: S. Kuremoto).

In this study, we built a bigger dataset for total single characters of Sen-O-So Guqin notation. 55 kinds of images of single characters were collected from 23 kinds of original Jianzi Pu from the Internet, and 4,951 single characters were obtained including data augmentation by image processing such as rotation, reversion, size enlargement (zoom in), size reduction (zoom out), and filtering. To realize online recognition, deep learning model YOLOv5 [9] [10] was adopted into the system as a classifier. The validation accuracy of YOLOv5 after 300 training arrived at 88.47% which was higher than the case of VGG16 87.50%. Additionally, audio/video data of single characters of Sen-O-So were created by Japan Society for the promotion of Guqin [2], and they were matched to the recognition results by YOLOv5.

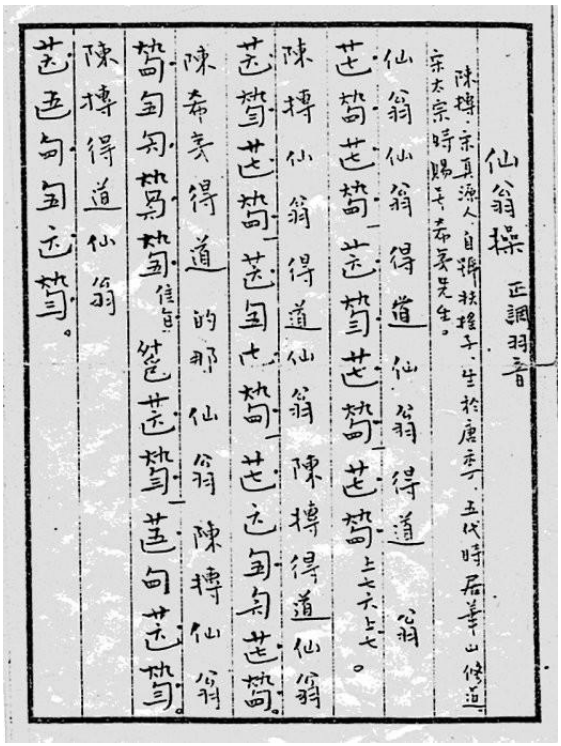


Fig. 3 A sample of Guqin musical notation “Sen-O-So” (https://www.sohu.com/a/155469920_713703).

Fig. 4 A sample of a single Jianzi Pu and its meaning in Japanese.

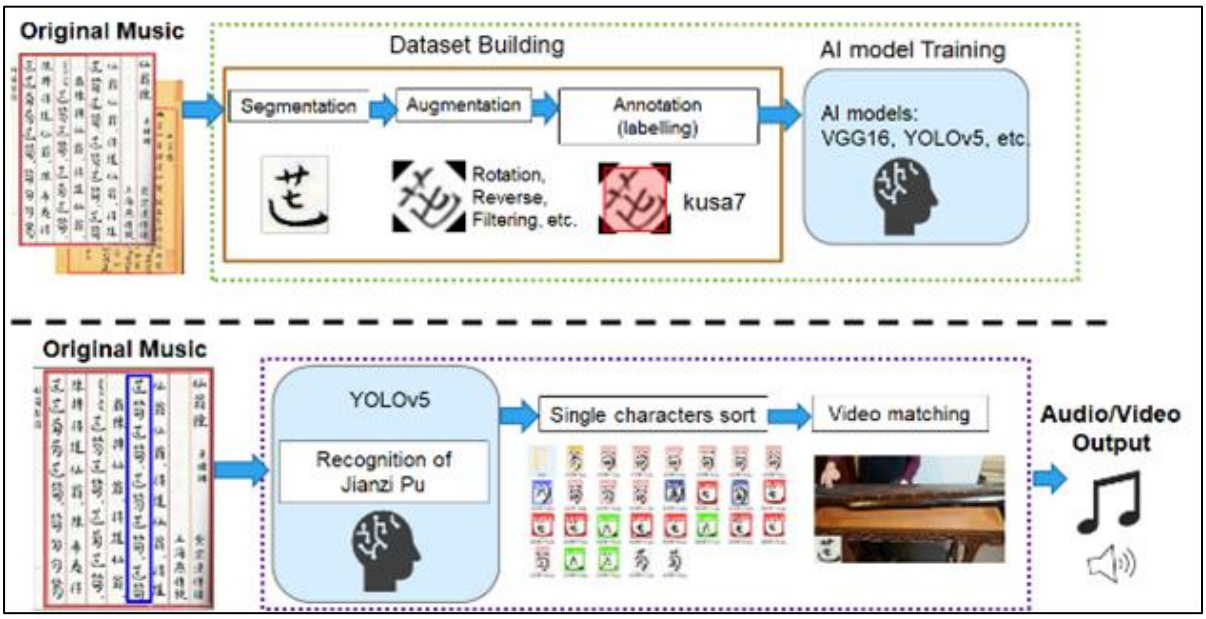


Fig. 5 A restoration system of Guqin music.

2. A Guqin Music Restoration System

A Guqin music restoration system, as shown in Fig. 5, is proposed in this study. It is aimed to realize that when a Jianzi Pu image is presented to the system, Guqin music in audio/video forms are output. The system is constructed by a fine-tuned deep learning model YOLOv5 [9] [10] and a matching process of the output of YOLOv5 and a single character database in audio and video forms.

2.1. *A Database of Jianzi Pu*

To train the deep learning models, a database of single characters of a well-known Guqin music "Sen-O-So" in Jianzi Pu form is built at first. 23 versions of Sen-O-So Jianzi Pu images were collected by the Internet. 55 kinds of single characters were segmented from the Sen-O-So notation, and totally 343 images of single characters were obtained. Data augmentation is adopted by image processing such as rotation, reflection, parallel translation, enlarge (zoom-in), reduce (zoom-out), filtering to the original single character images. As the result, 4,951 images of single characters were obtained as a database of Jianzi Pu. A sample of the original single character images and a sample of data augmentation were shown in Fig. 6 and Fig. 7, respectively. In Fig. 6, 55 single characters of Sen-O-So are listed as a sample of 23 versions of the music collected from the Internet. In Fig.7, the original single character and its variations obtained from image processing are listed.

Fig. 6 A sample of 55 single characters of Sen-O-So.

2.2. *Deep Learning Models*

A well-known deep learning model VGG16 [7] [8] and an online object recognition model YOLOv5 [9] [10] were investigated their recognition accuracies to single characters of Jianzi Pu. The state of art of YOLO is YOLOv8, and we are investigating its performance for Jianzi Pu recognition recently.

3. Experiments and Results

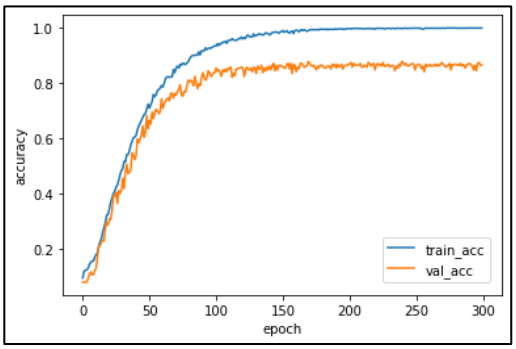
Using the database of Sen-O-So described in Section 2, VGG16 and YOLOv5 were fine-tuned with 300 epochs. The average recognition accuracies of the single characters were 87.50% and 88.47% respectively. The performance of YOLOv5 stood its priority to VGG16 for its more complicated backbone (CSP Darknet53) which

serves higher feature extraction ability comparing to VGG16 which is a classic deep learning models composed of convolutional neural networks (CNN) and multi-layered perceptron (MLP). So YOLOv5 was adopted as the classifier in our line restoration system.

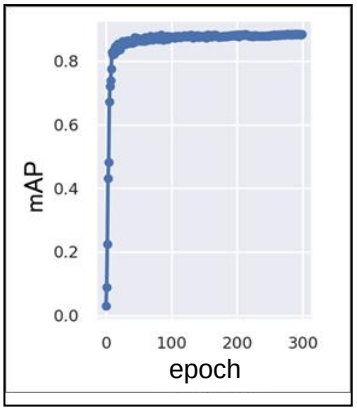
ファイル名	画像	画像処理の詳細
7 (14).jpg		元の画像
7 (14).jpg_**30**.jpg		高さと幅を30pxに変更
7 (14).jpg_**100**.jpg		高さと幅を100pxに変更
7(14).jpg_**200**.jpg		高さと幅を200pxに変更
7 (14).jpg_**blur**.jpg		カーネルサイズ（7，7）標準偏差1.5で平滑化
7 (14).jpg_**brighter**.jpg		色の値を1.5倍
7 (14).jpg_**darker**.jpg		色の値を0.9倍
7 (14).jpg_**noise**.jpg		ランダムな場所を対象にRGBランダムに選んでその値をランダムな少数分±する
7 (14).jpg_**r45**.jpg		反時計回りに45度回転
7 (14).jpg_**r90**.jpg		反時計回りに90度回転
7 (14).jpg_**r135**.jpg		反時計回りに135度回転
7 (14).jpg_**r180**.jpg		反時計回りに180度回転
7 (14).jpg_**r225**.jpg		反時計回りに225度回転
7 (14).jpg_**r270**.jpg		反時計回りに270度回転
7 (14).jpg_**r315**.jpg		反時計回りに315度回転

Fig. 7 A sample of data augmentation.

The change of accuracies in training of two models is shown in Fig. 8. It can be confirmed that the accuracies (training accuracy and validation accuracy) of VGG16 were converged from 200 epoch (Fig. 8a), meanwhile, it converged from 100 epoch in the case of YOLOv5 (Fig. 8b).



a. VGG16



b. YOLOv5

Fig. 8 Learning curves of VGG16 and YOLOv5.

Fig.9 A sample of audio/video data matched to a single character of Jianzi Pu.

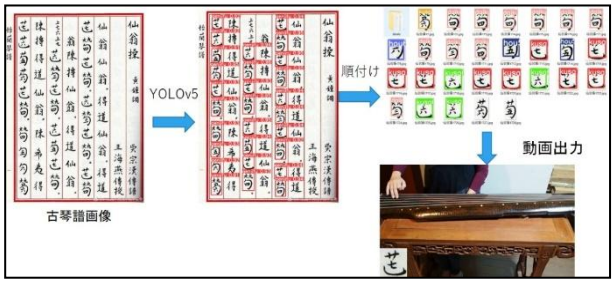


Fig.10 A sample of the processing result of the proposed Guqin music restoration system.

Single characters of Jianzi Pu were played by a professional Guqin player, Mrs. Yuki Takei [2] and recorded in audio/video forms as shown in Fig. 9. The order of single characters in a Guqin musical notion was given by their position detected by YOLOv5. By matching the recognized single characters to audio/video data, the system output music and videos as shown in Fig. 10. In Fig. 10, the input image was captured by a webcam, and then YOLOv5 detected Jianzi Pu single characters as an object recognition task recording the position of each character, finally, the detected single characters were matched to the video clips and the output of the system was the continuous plays of those matched videos.

4. Conclusion

Guqin (古琴, Chinese seven-string zither) music notation written in Jianzi Pu style, were recognized by deep learning methods such as VGG16 and YOLOv5. The main contributions of this study include following: 1) We created a dataset of a well-known Guqin music Sen-O-So (仙翁操, Melody of the Immortal Elder) of single characters which includes 55 kinds of single characters of Sen-O-So in 4,951 images from 23 versions found in the Internet and obtained by data augmentation; 2)We created a dataset of Sen-O-So video clips to match the recognition results of single characters by YOLOv5; 3)An online restoration system was developed by YOLOv5 which realized the transformation of Jian-Zi-Pu of Sen-O-So to audio-visual content that suggested the restoration of the Guqin music was executable. To verify the performance of YOLOv5 compared to other deep learning models, VGG16 was investigated in the experiments.

To restore music from Guqin notation (Jianzi Pu) through Dapu, recognizing the Jianzi Pu symbols is only a starting point; it alone does not allow full musical restoration. [10] This is because rhythm, one of the three fundamental elements of music, is absent from Jianzi Pu. Adding rhythm requires a deep understanding of the musical intent behind each piece. To grasp this intent, interpretations of the piece's title and explanations provided in Classical Chinese are crucial. By applying natural language processing (NLP) techniques, we can generate textual expressions that convey the intent of each piece. This generated text can then be aligned with the music by training on sample data that matches text with musical elements.

The Dapu process used by experts can serve as a foundation for designing algorithms to automate this restoration using AI. By analyzing the steps experts follow in interpreting and reconstructing music from Guqin tablature, these steps can be modeled and implemented in AI systems, allowing for a structured approach to automatic music restoration and we leave these works in the future.

Acknowledgements

This work was supported by JSPS KAKENHI Grant (No. 22H03709, No.22K12152)

References

1. China Script, "The difference between Qin and So", http://chugokugo-script.net/chugoku-bunka/kin.html (2016) (in Japanese)
2. Japan Society for Promotion of Guqin: https://www.guqin.jp/about
3. R. Takao, "A survey and analysis of the Dapu process in Qin music", Journal of the Musicological Society of Ochanomizu University, No. 10, 2008 pp.40-45 (in Japanese)
4. Z.X. Pan, C. L. Zhou: "Text Segmentation and Extraction from Images of Guqin Jianzi Pu", Mind and computation, Vol.1, No.2, pp. 281-290 (2007) (in Chinese)
5. Chen Shi: "Guqin Notation and Music Style Recognition", Computer Science, Corpus ID: 31634157 (2016)
6. L. Wang, Y. Sun, Z.L. Luo, H. Zhang: "AI Automatic Translation of Jianzi Pu A Case Study of "Stream" and "Non-Staining", Art Education, Vol.3 (2019) (in Chinese)
7. B. Yang, M. Sato, S. Kuremoto, M. Koshiba, S. Mabu, H. Hieda, T. Kuremoto, "Recognition of Guqin music by deep learning methods", Proceedings of the Electronics, Information and Systems Conference Electronics, Information and Systems Society, I.E.E. of Japan, pp.512-515 (2022) (in Japanese)
8. Bowen Yang, Shun Kuremoto, Mamiko Koshiba, Shingo Mabu, Hiroo Hieda, Takashi Kuremoto, "Recognition of Guqin Notation using Deep Learning", in Proceedings of Innovative Application Research and Education (ICIARE2022), pp.61-64 (2022)
9. J. Redmon, S. Divvala, R. Girshick, A. Farhadi, You Only Look Once: Unified, Real-Time Object Detection, aiZiv:1506.02640v5 (2016)
10. T. Kuremoto, "Guqing Music Recognition by Machine Learning Methods", Impact, pp.40-42, 2024(1), DOI: 10.21820/23987073.2024.1.40 (2024)

Authors Introduction

Prof. Dr. Takashi Kuremoto

He received his B.E. degree from the University of Shanghai for Science and Technology, China, in 1986, M.E. and Ph.D. degrees from Yamaguchi University, Japan, in 1996 and 2014, respectively. He worked as a system engineer at the Research Institute of Automatic Machine of Beijing from 1986 to 1992 and until 2021, he was an Assistant Professor at Yamaguchi University. He was an Academic Visitor of the School of Computer Science, the University of Manchester, U.K., in 2008. Currently, he is a Full Professor with the Department of Information Technology and Media Design, Nippon Institute of Technology, Japan. His research interests include artificial neural networks, fuzzy, evolutionary computation, time series forecasting, and intelligent systems

Mr. Kazuma Fujino

He received his B.E. degree from Nippon Institute of Technology, Japan, in 2023. He works at SET Software Co., Ltd., Tokyo, Japan, currently.

Mr. Hirokazu Takahashi

He received his B.E. degree from Nippon Institute of Technology, Japan, in 2023.

Mr. Shun Kuremoto

He received his B. F. degree from Shanghai University of Finance and Economics, China, in 1989, and M. A. in Commerce degree from Waseda University, Japan, in 1997. Currently, he is a guest researcher of the Institute for Future Engineering, Japan, as well as Director of Japan Society of Promotions of Guqin, and a doctoral program student in Yamaguchi University, Japan.

Dr. Mamiko Koshiba

She received her B.E., M.E. and Ph.D. degrees from Tokyo University of Agriculture and Technology, Japan, in 2001, 2002 and 2005, respectively. She worked as a postdoctoral neuroscientist at National Center of Neurology and Psychiatry 2005 to 2007 and until 2013, she was a Visiting Associate Professor at Tokyo University of Agriculture and Technology and from 2013 to 2016 at Saitama Medical University, Currently, she was an Associate Professor of Yamaguchi University, Japan. Her research interests include applications of artificial neural networks for neuronal system evolution and development.

Mr. Hiroo Hieda (1945-2023)

He was a Fellow of the Institute for Future Engineering. Longtime, he did comprehensive research on science and art. Currently, he researches on the history of Qin (七絃琴) in early modern Japan, especially in the Edo period. He published a book named 近世琴学史攷(A study to the history of Qin in early modern Japan) in 2018.

Prof. Dr. Shingo Mabu

He received his B.E and M.E. degrees from Kyushu University, Japan, in 2001 and 2003, respectively, and his Ph.D. degree from Waseda University, Japan, in 2006. From 2006 to 2007, he was a visiting lecturer at Waseda University, Japan, and from 2007 to 2012, an assistant professor at the same university. From 2012 to 2017, he was an assistant professor at Yamaguchi University, Japan, and from 2017 to 2020, an associate professor at the same university. He is currently a professor at the Graduate School of Sciences and Technology for Innovation at the same university.