2024 IEEE International Conference on Robotics and Automation (ICRA 2024)
May 13-17, 2024. Yokohama, Japan

X-Tacformer : Spatio-tempral Attention Model for Tactile Recognition

Jiarui Hu, Yanmin Zhou*, Zhipeng Wang, Xin Li, Yongkang Jiang , Bin He

Abstract— **Recently, tactile sensing has attracted great interests in robotics, especially for exploring unstructured objects. Sensor arrays play an important role in the exploration, which generates rich spatio-temporal information. In this work, we propose an efficient tactile recognition model, X-Tacformer. This model pays attention to both spatial and temporal features of tactile sequences from sensor arrays, which is verified by four public datasets, Ev-Objects, Ev-Containers, Augment8000 and BioTac-Dos. Comparative studies show that our model has resulted in a significant improvement of the recognition accuracy by 0.0223, 0.1416, 0.2735 and 0.1592 in these datasets. In order to verify its performances on dataset with rich spatio-temporal features, a self-designed dataset, ALU-Textures, was constructed with 10 fabrics from everyday textiles, aiming to extend the data collection action modes of current datasets by simulating human rubbing movements with the thumb and index fingers of an Allegro hand. Our model also demonstrates efficient salient feature learning capabilities on ALU-Textures, which is further augmented by tactile data augmentation methods.**

I. Introduction

It is well known that the sense of touch is one of the most important information sources for both humans and robots to perceive the object properties. With the obtained tactile information, human can easily judge the type of an object [1]. When we use hands to grip an object, the information of object's hardness, texture and weight provides important cues for fine object recognition [2]. In the dynamic exploratory procedure, the contact surface is a key to acquire useful information. Sensor arrays with a higher spatial density units can provide rich contact surface information from different objects [3]. The tactile data collected by a sensor array at each time step can be viewed as a tactile image, containing spatial information among sensor units. Meanwhile, sequence of tactile images would be generated during the dynamic exploratory process of objects.

These series of tactile images carry temporal information of tactile perception. These tactile data with rich spatio-temporal features are beneficial for object explorations.

In this work, we combine the temporal data processing module with the spatial data processing module to form X-Tacformer, which extracts multiple dimensional tactile features by both spatial and temporal attentions to extract multiple taxels tactile features. The X-Tacformer is adopted from Transformer [4]. As a neural network, its performance relays heavily on the characteristics and size of the datasets. Therefore, four public datasets, Ev-Objects [5], Ev-Containers [5], Augment8000 [6] and BioTac-Dos [7] are chosen to validate the performances of the X-Tacformer. Since the spatio-temporal features of these datasets are limited, a high-density arrayed tactile dataset, ALU-Textures is constructed by using robotic hand to simulate human tactile recognition action of rubbing. After touching the objective fabric, the two fingertips move in opposite directions, producing relative motions. The densely distributed 6×6 u-skin [8] sensor array on each fingertips is activated during these processes. The tactile data is collected on each taxel on these two sensor arrays at each time step.

The authors are all with the Department of Control Science and Engineering, College of Electronics and Information Engineering, Tongji University, Shanghai 201804, China, with National Key Laboratory of Autonomous Intelligent Unmanned Systems, Shanghai 201210, China, and Frontiers Science Center for Intelligent Autonomous Systems, Shanghai 201210, China.

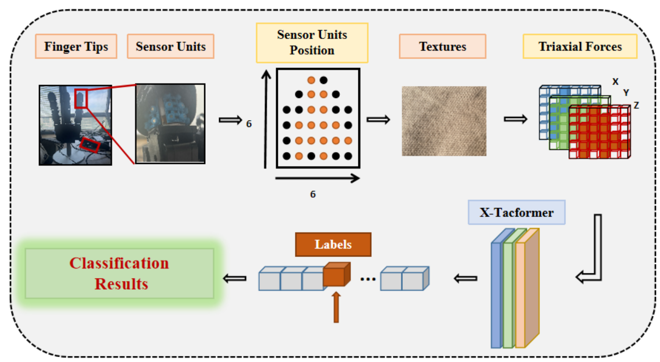

Fig. 1. An overview of the texture classification using tactile data. From the left side, two fingertips with sensor arrays receiving tactile stimulations. The orange points stand for the activated sensor units. The triaxial forces are sensed by the units of the sensor arrays. The tactile data are input into X-Tacformer.

The process for data access and utilization is shown in Fig.1. We also perform an expansion operation on the dataset to improve the model performance on small samples which has not been widely studied with tactile datasets. The performance verification of X-Tacformer is conducted on the above five datasets. The results show that our X-Tacformer boosts the recognition accuracy to a large extent.

The main contributions of this paper include:

- We develop a spatio-temporal attention model, X-Tacformer, which take care of salient features in tactile recognition.
- We use Allegro hand equipped with u-skins to collect a tactile dataset with 200 samples on 10 fabrics. We further explore the tactile data augmentation methods on tactile dataset.
- A set of experiments demonstrate our proposed method improves the tactile recognition accuracy significantly.

979-8-3503-8457-4/24/$31.00 ©2024 IEEE

II. Related Works

Object recognition is one of the core capabilities of intelligent robots. It involves inferring the category of the objects with perception data.

A. Tactile Recognition

The tactile recognition often requires processing the data to a specific format that can be applied to models, such as tactile image sequences, tokens, adjacency matrix to facilitate the feature extractions, such as spatial and temporal features.

Constructing tactile data into image sequences is a common pre-processing method. Sarakona et al. [9] used tactile images as inputs to train a neural network. They added an MLP classifier to a pretrained VGG-16 model. Yuan et al. [10] used Gelsight to acquire tactile images to achieve fabric Classification. Similarly, Cao et al. [11] used Gelsight to generate tactile image inputs to validate their spatio-temporal attention model for tactile texture recognition. They all fed tactile image sequences into image processing models. These models based on image processing methods could learn the difference between tactile images at different time step as well as acquire the spatial feature of taxels. Transformer has shown to achieve remarkable success in the field of image processing [4], [12].

Graph neural network has been applied to extract the spatial features of tactile information. Garcia et al. [13] used Non-matrix tactile sensor information to predict the stability of grip. They described the actual position of units on sensor array by graph. Gu et al. [5] proposed a tactile object recognition method based on event-driven tactile sensor arrays. They manually constructed a tactile graph based on the spatial location of the sensors. They used Graph Convolutional Networks(GCN) and Spiking Neural Networks(SNN) to build their model and tested the model on two datasets. The integration of sensor units in non-matrix form through GCN enables individual units to learn some features from neighbors.

Due to the time dependance nature of tactile sequences, several studies have combined time-series processing model with CNN in order to obtain both of the spatial and temporal features from tactile data. Li et al. [14] used CNN to concatenated features to create mixed feature vectors and passed the mixed vectors in sequences to LSTM. This model, which was stitched together in sequence, was named CNN-LSTM. Zhang et al. [15] presented an applications of slip detection using ConvLstm, a special LSTM with internal convolutions to learn spatio-temporal tactile features. They viewed a series of values at each sensing unit as a token and introduced LSTM to learn the temporal properties in tactile information. The mixed thought has yielded progressive results. The combination of GCN and GRU is hope to get more fully spatial and temporal representations of tactile information [16].

B. Tactile Data Acquirement and Augmentation

Four public tactile datasets, Ev-Objects [5], Ev-Containers [5], Augment8000 [6] and BioTac-Dos [7] are often used for the studies of object classification. The Ev-Objects comprises tactile data from 36 object classes. They were collected by a robot gripper equipped with NeuTouch sensor array during the process of grasping lifting objects off the table by 20 cm during 5s. The Ev-Containers includes tactile data collected with the same approach of Ev-Objects but on four containers. Both of these two datasets contain dynamic sequences data with spatio-temporal features. The Augment8000 was collected by a three-fingered gripper with a 3D force sensor by performing an active tactile glance on 20 objects. The BioTac-Dos was generated by slipping 11 objects on a BioTacSP sensor at 7 directions of slip: slip north, slip west, slip south, slip east, rotate clockwise, rotateanti-clockwise and stable contact. 11 different objects were slipped on a BioTacSP sensor. The dynamic data generated by the sliding process contains rich spatio-temporal features.

The way of data acquirement affects the features of the data. Lots of fabric classification tasks used one-directional movements to actively sense fabrics by tactile perception. Taunyazov et al. [17] fixed a tactile sensor on the end of a robot arm to contact a fabric. Then the robot arm was controlled to drag the sensor in one direction to sense the fabric. Similary, Gupta et al. [18] used a tactile sensor on an end-effector of a robotic arm to collected the tactile information during the one-directional slidings on fabrics under 1 N normal force at three sliding speeds. However, the bi-directional rubbing movement for tactile fabric exploration is infrequently studied, which is commonly used by humans. Tactile dataset collected by such movement is absence as well.

Data augmentation methods benefit the size and quality of datasets for nural network based studies. Current research on tactile data augmentation is relatively scarce. Image enhancements in visual perception cannot be directly applied to tactile data due to their very different perception mechanisms. Guennec et al. [19] warped randomly selected time-series of tactile data for tactile data augmentation. Such method not only could get more samples but also ensured high usage of the data. T. Um et al. [20] applied several data augment techniques, such as scaling, magnitude-warping and jitter to their tactile dataset and achieved positive effects in their study.

III. The X-Tacformer Construction

In this section, we ill present our design of X-Tacformer, as illustrated in Fig. 2, which consists of two branches: temporal graph attention module and spatial graph attention module.

A. Spatial Graph Attention Module

In order to get the spatial information from each tactile image, we construct a spatial graph attention module(Tacformer). This branch gives different attentional weights on different taxels, enabling fuller learning of spatial information. Inspired by [5], we employ TAGConv in this module to better match the topology of the tactile graph [21].

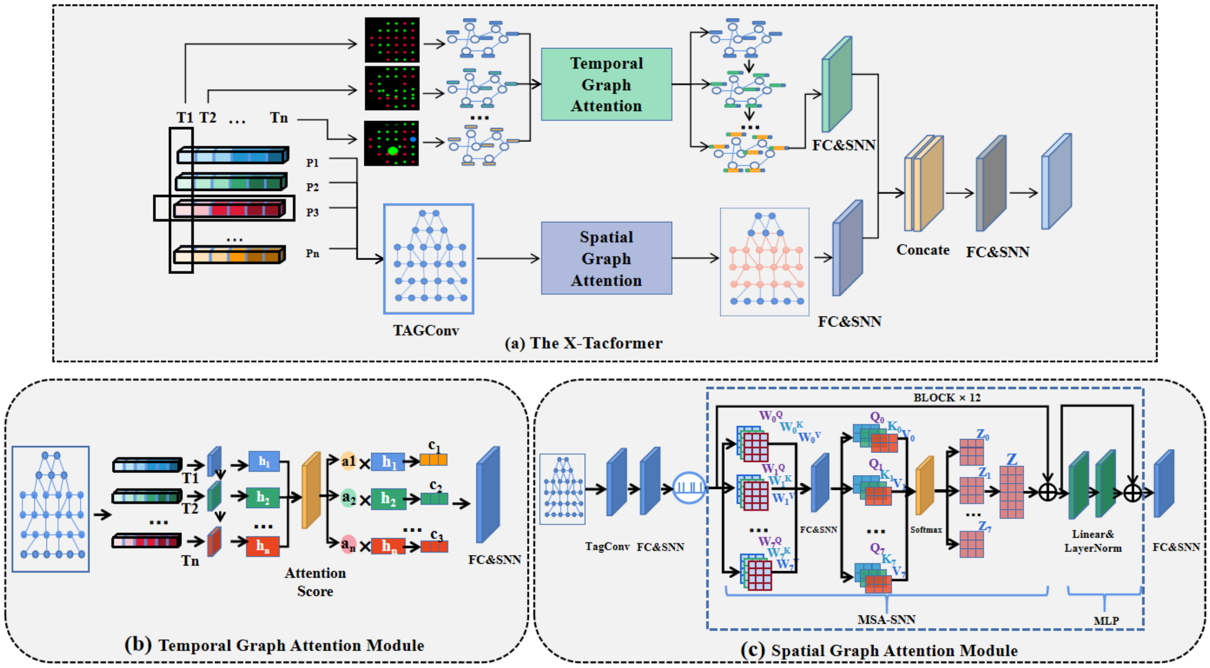

Fig. 2. The X-Tacformer Framework. (a) The X-Tacformer are divided into two branches (b) The Temporal Graph Attention Module, which processes a series of tactile images, and (c) The Spatial Graph Attention Module, which processes a tactile graph.

A TAGConv operation is defined as:

$$H_l = \sum_{c=1}^{C} G_{c,l} * h_c + b_l \tag{1}$$

H is the output feature map of layer l. h_c is the input feature of node c. C is the number of input features of each node. b_l is a learnable bias vector. $G_{c,l}$ is the convolution kernel.

The integrated feature of a tactile image is fed into a Multi-head Self-attention(MSA-SNN) functional unit. We use SNN to fine tune our model [22]. As the basic unit of the SNN, LIF receives stimuli and compares the accumulated membrane potential to a threshold value. It determines whether a pulse should be generated [23]. Here is the update of the membrane potential description:

$$H(t) = V(t-1) + \frac{1}{\tau}[X(t) - (V(t-1) - V_{\text{reset}})] \tag{2}$$

$$S(t) = \theta[H(t) - V_th] \tag{3}$$

$$V(t) = H(t)[1 - S(t)] + V_{reset}S(t) \tag{4}$$

where τ is the membrane time constant and $X(t)$ is the input current at time step t. LIF will trigger a pulse when the membrane potential $H(t)$ exceeds the trigger threshold. S(t) is the step function, which equals to 1 when $\theta(t) \geq 0$. $V(t)$ denotes the membrane potential triggered by events.

The MSA has three floating-point key components, namely $query$, key and $value$ which learn from the same source [24].

$$Q = XW, K = XW, V = XW \tag{5}$$

The output of self-attention can be computed as Equation6:

$$MSA(Q, K, V) = Soft\max(QK^T)V \tag{6}$$

However, the calculation of MSA is not applicable in SNN for two reasons: all of the matrix multiplication of Q, K and softmax function do not comply with the calculation rules of SNN. The quadratic space and time complexity of the sequence do not meet the efficient computational requirements of SNN. To address these problems, a linear layer is added. As shown in Fig.2, the MSA-SNN is defined as:

$$Q, K, V = LN(Linear(LIF(LN(Linear(X, W))))) \tag{7}$$

$$MSA'(Q, K, V) = Soft\max(\frac{QK^T}{\sqrt{dK}})V \tag{8}$$

$$MSA = LN(Linear(LIF(MSA'))) \tag{9}$$

B. Temporal Graph Attention Module

Dynamic tactile exploration process by a tactile sensor array generates sequences of tactile images. In order to obtain the temporal information from such tactile data, we construct a temporal graph attention module. This T-GCN could not only organize data of multiple taxels from a tactile array but also capture the changes in dynamic characteristics of tactile sequences. The feature of all taxels on the same time step is put into the GCN to construct a tactile image. The tactile images are input into GRU frame by frame. The GCN and GRU are combined by these equations:

$$u_t = \sigma(W_u * [GC(A, X_t), h_{t-1}] + b_u) \tag{10}$$

$$r_t = \sigma(W_r * [GC(A, X_t), h_{t-1}] + b_r) \tag{11}$$

$$c_t = \tanh(W_c * [GC(A, X_t), (r_t * h_{t-1})] + b_c) \tag{12}$$

$$h_t = u_t * h_{t-1} + (1 - u_t) * c_t \tag{13}$$

where GC stands for graph convolution process. c is the current stored content. h_t is the output state at moment t.

The Temporal Graph Attention Module is shown in Fig.2 The attention mechanism could adjust the importance of frames.

IV. The Dataset Construction

This section provides details of the procedures of ALU-Textures collection. The data acquisition process is built on Robot Operating System (ROS) of a Linux machine.

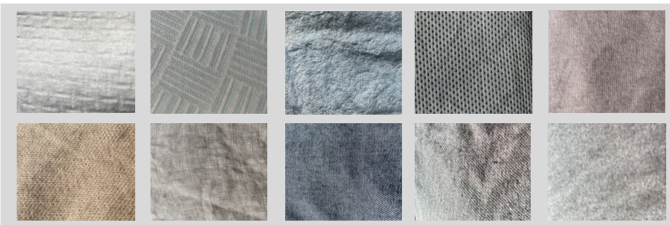

Fig. 3. Fabrics used for our dataset. From left to right, top to bottom: washcloth, pillowcase, pillow cloth, synthetic fiber T-shirt, cotton T-shirt, cotton shorts, sackcloth, denim, cotton and linen pants, blazer pants.

A. Allegro Hand with Tactile Sensors

The working environment for the Allegro hand is the ROS-Hubmle, runing on Ubuntu-22.04. Two ESD-CAN BUS could build communication between the PC and the controller of both hand and the u-skin driver. The Allegro hand has four fingers, each finger has four degrees of freedom. Every fingertip is covered with a u-skin sensor. The u-skin sensor is a type of high-density sensor array which contains 30 sensor units, as shown in Fig.1. It could collect forces in three axes.

B. Input Data Acquirement

Two robotic fingers are controlled to rub 10 fabrics bi-directionally. Initially, the fingertips of the thumb and index finger touch fabrics to activate the tactile sensors. Then the rotation angles of the roots of these two fingers are controlled to change to provide sufficient friction on fabtics. When the root of the index finger root rotates 5 degrees, that of the thumb finger root rotates -5 degrees, producing relative motion for 5 seconds. Finger trajectories can be viewed in rviz. The movement process is showing in Fig.1.

The tactile signal starts to be recorded once the fingers start to move. The initial sensor value is set to 0. Whenever a unit of a sensor array touches the fabric, it is activated and its reading changes. An example reading change of a unit can be seed in Fig.1. The offset of the green dot is the reading of the tangential forces.

10 common fabrics in our daily life are chosen to construct the ALU-Textures dataset, as shown in Fig.3. For each of the 10 objects, 20 experimental trials are per-formed resulting in 200 samples in total. At the same time, the joint information throughout the series of finger rubbing is recorded. Thus, the ALU-Textures has a total of 2 (u-skin sensors) × 30 (taxels) × 3 (triaxial forces) + 8 (joint angles) = 188 measurements. When two fingertips touch the fabrics, 50 time steps value starts to be saved. These tactile data is reformed to $X = [C, D, T]$ as the inputs. Fig.4 depicts the acquisition and storage procedures of the tactile data.

Augmentation operations on our small sample dataset are performed. In order to protect the temporal and spatial integrity of the data, random cropping is not considered in this work. Therefore, we adopt the method proposed by T. Um et al. [20] to extend ALU-Textures. The basic ideas of these four methods are: 1) Jitter: A random Gaussian noise with zero mean and a standard deviation of 0.05 is added to the dataset. The important features can be made more salient. 2) Scaling: The time-series data for every feature is multi-plied with a random factor, the mean value of 1 and the standard deviation of 0.1, respectively. This operation equals applying constant noise to the samples. 3) Magnitude Warping: Each feature time serie data is convolved with a smooth random curve created with the cubic spline method with 4 knots around the value of 1 with a standard derivation of 0.05. All these three of these methods process the values at each time step for feature enhancement. The fourth method processes the data time in intervals. 4) Time warping: The time intervals between the samples are smoothly distorted, thereby generating signal shifts at their locations. We use these methods to bring variances to data values at each time step, without affecting the number and length of time steps. These augment methods can increase the sample sizes while maintaining the data integrity. Then the joint angles can be accordingly calculated and associated to each augmented tactile sequence.

V. Experiments

In our experiments, we aim to validate whether the X-Tacformer can improve the performance of tactile recognition and explore the effectiveness of our augmentation methods for tactile data. Two studies are conducted: 1) The accuracy performance of X-Tacformer is verified by comparative studies with baseline models on public datasets. 2) Four augmentation methods are used to extend ALU-Textures. The ALU-Textures and aug-ALU-Textures are tested by spatio-temporal models to examine the performance of augmentation approaches.

A. The Model Performance Experiments

By comparing the recognition accuracy with basel spatio-temporal models, we are able to more intuitively explore the effect of X-Tacformer.

It can be seen from TABLE.I and TABLE.II, we can see that our proposed X-Tacformer achieves the best recognition performance on all datasets. The SGnet was reported the best model for Ev-Objects and Ev-containers in their study [5]. Compared with the SGnet, our X-Tacformer model has increased the accuracy by 0.0223 and 0.1416 on these two datasets, respectively. The D-CNN performed the best on Augment8000 in their study [6]. Compared with the DCNN, our X-Tacformer model has increased the accuracy by 0.2735. The ConvLSTM was shown as the best model for

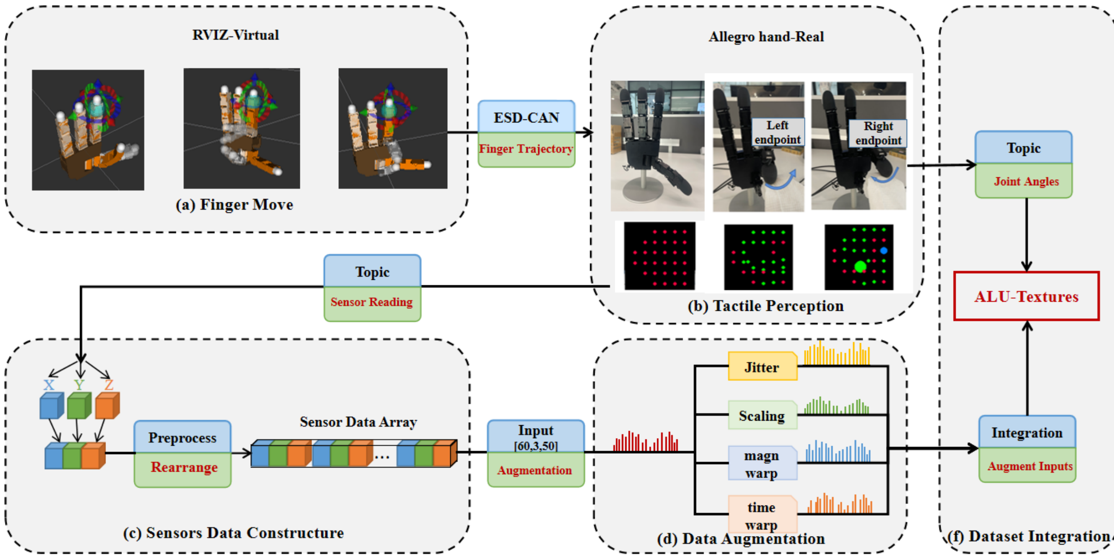

Fig. 4. Data Acquisition Framework. (a) The finger_move program controls the eight joints of the thumb and index fingers. The ESD-CAN converter establishes the connection with the real Allegro hand. (b) The trajectory of the Allegro hand. (c) The sensor readings are passed by the CAN-BUS and formed given format. (d) The four augmentation methods are used to expand ALU-Textures. (f) The joint angles are concatenated with the tactile data, as well as the labels are marked.

BioTac-DoS in their study [7]. The BioTac-DoS has three subcategories: Solids set, Small objects set and Textures set. Compared with the ConvLSTM, our model has increased the accuracy on all subcategories. The accuracy is increased by 0.1592, 0.0568 and 0.1453 on the sets of Solids, Small object and Textures. On our own dataset, the X-Tacformer has achieved the highest accuracy. These experiments show that X-Tacformer has better spatio-temporal feature learning properties.

TABLE I

THE ACCURACY ON FOUR PUBLIC DATASETS WITH DIFFERENT MODELS

Model	Ev-objects	Ev-contains	Augment8000	BioTacSP
SGnet	0.8944	0.6417	-	-
D-CNN	-	-	0.7258	-
ConvLSTM	0.8611	0.7333	0.7008	0.7380
CNN-LSTM	0.8472	0.7167	0.7208	0.8182
Tacformer	0.9027	0.7500	0.9650	0.8442
X-Tacformer	**0.9167**	**0.7833**	**0.9993**	**0.8833**

TABLE II

THE ACCURACY OF ALU-TEXTURES AND AUG-ALU-TEXTURES

Model	ALU-Textures	aug-ALU-Textures
ConvLSTM	0.7426	0.8273
CNN-LSTM	0.7460	0.8133
Tacformer	0.7978	0.8333
X-Tacformer	0.8776	**0.9295**

B. The Augmentation Experiments

There is a close correlation between the learning ability of a neural network and the size and quality of the dataset [25]. In order to further improve the performance of our spatio-temporal model. Four suitable augmented methods are used on our dataset. TABLEII displays the accuracy of different spatio-temporal models on our dataset. The accuracy of aug-ALU-Textures has increased by 0.0519 with X-Tacformer. This dataset expansion results in an increasing performance with these data. The friction generated by rubbing on fabrics is used as an important classification indicator. The friction generated by similar materials has influence in classification. As shown in Fig.5, the sixth and ninth textures, cotton fabrics, have strong material similarities. The fifth and tenth fabrics have the akin touch sensation, leading to reduced categorization effectiveness.

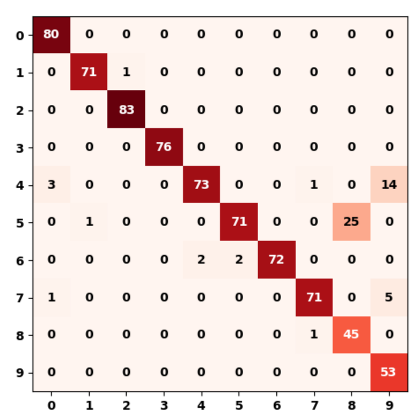

Fig. 5. The recognition of X-Tacformer on aug-ALU-Textures. The training labels is corresponding to the labels in Fig.3

These experiments have demonstrated that the time-series tactile data could be augmented without changing data structure to improve the model training effects effectively.

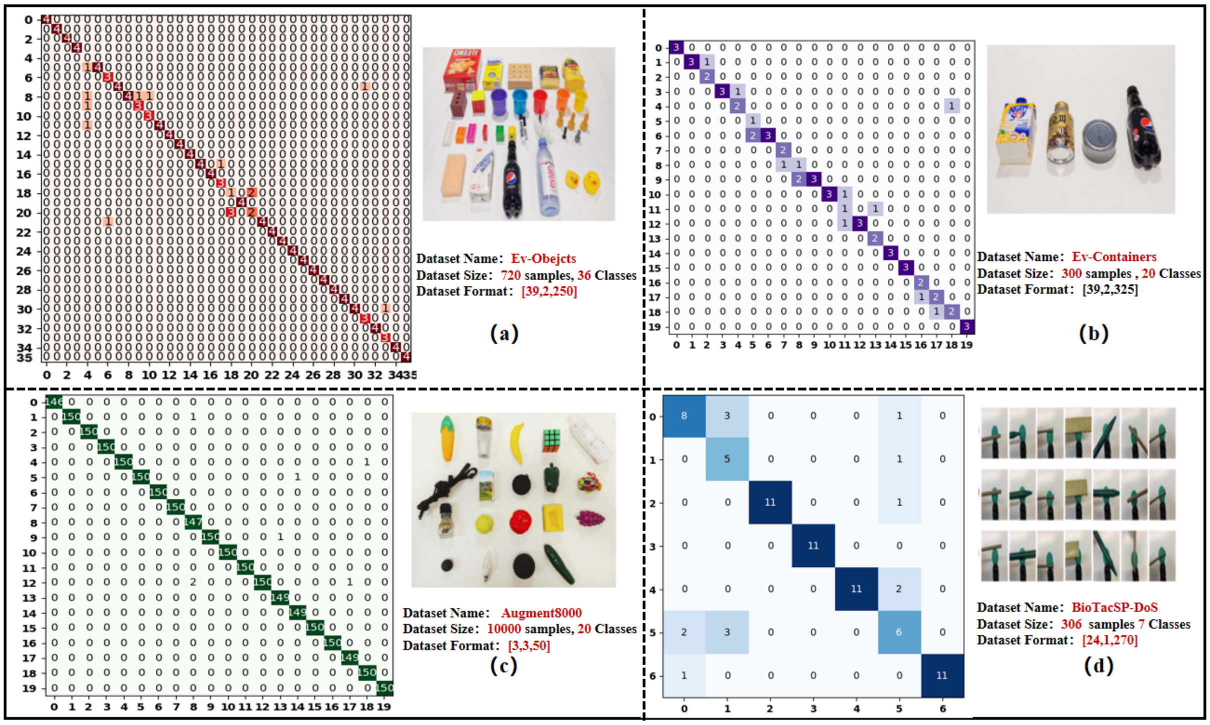

Fig. 6. The recognition results of X-Tacformer on four public datasets.

VI. Conclusions

In this work, we propose a spatio-temporal attention model, X-Tacformer, for tactile recognition, which not only pays attention to spatial feature of tactile image but also temporal feature of the sequences of tactile images. The spatial attention module could learn the tactile features of units on a sensor array, which are generated by their graph structure by perceiving objective objects. The temporal attention module could acquire the tactile information from the time sequences of tactile images. The X-Tacformer can fully utilize the spatio-temporal properties of tactile data for object recogniton.

The Allegro hand is fitted with 3D force sensors on each fingertips. It performs an active tactile rubbing action on a set of 10 objects to construct the dataset of ALU-Textures. The tactile data augmentation methods, Jitter, Scaling, Magnitude Warping and Time warping, are studied to augment the ALU-Textures. The baseline neural network models, CNNLSTM, ConvLSTM, Tacformer and X-Tacformer are trained on Ev-Objects, Ev-Containers, Augment8000, BioTacSP-Dos, and ALU- Textures.

Results show that, the spatially and temporally selective attention mechanism of our model have resulted in a significant improvement of the recognition accuracy. The X-Tacformer is able to select informative features efficiently with small sample data at both space and time dimensions. The augmentation methods result in a significant model performance increase on small sample data by maintaining the original data structure. Our proposed X-Tacformer could be applied to facilitate robotic tasks with the spatio-temporal tactile perception requirements, such as grasping and object sorting.

VII. Acknowledgement

This work was supported in part by the National Key Research and Development Program of China (No. 2020AAA0108905), in part by the National Natural Science Foundation of China (No. U2013602, 62088101, 61825303), in part by the Science and Technology Commission of Shanghai Municipality (No. 2021SHZDZX0100, 22ZR1467100, 23YF1449700), and the Fundamental Research Funds for the Central Universities. (Corresponding author: Yanmin Zhou, email: yanmin.zhou@tongji.edu.cn.)

References

[1] Das A, Alagirusamy R. Improving tactile comfort in fabrics and clothing[M]//Improving Comfort in Clothing. Woodhead Publishing, 2011: 216-244.

[2] Bensmaia S, Hollins M. Pacinian representations of fine surface texture[J]. Perception & psychophysics, 2005, 67: 842-854.

[3] Gorges N, Navarro S E, Göger D, et al. Haptic object recognition using passive joints and haptic key features[C]//2010 IEEE International Conference on Robotics and Automation. IEEE, 2010: 2349-2355.

[4] Dosovitskiy, A., Beyer, L., Kolesnikov, A., Weissenborn, D., Zhai, X., Unterthiner, T., & Houlsby, N. (2010). An image is worth 16x16 words: Transformers for image recognition at scale. arXiv 2020. arXiv preprint arXiv:2010.11929.

[5] Gu F, Sng W, Taunyazov T, et al. Tactilesgnet: A spiking graph neural network for event-based tactile object recognition[C]//2020 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS). IEEE, 2020: 9876-9882.

[6] Maus P, Kim J, Nocentini O, et al. The Impact of Data Augmentation on Tactile-Based Object Classification Using Deep Learning Approach[J]. IEEE Sensors Journal, 2022, 22(14): 14574-14583.

[7] Zapata-Impata B S, Gil P, Torres F. Learning spatio temporal tactile features with a ConvLSTM for the direction of slip detection[J]. Sensors, 2019, 19(3): 523.

[8] Funabashi S, Yan G, Hongyi F, et al. Tactile transfer learning and object recognition with a multifingered hand using morphology specific convolutional neural networks[J]. IEEE Transactions on Neural Networks and Learning Systems, 2022.

[9] Sarakona K P, Sakaib Y, Serikawaa S. Object shape and force estimation using deep learning and optical tactile sensor[C]//Proc. 6th Inst. Ind. Appl. Eng. Int. Conf. Intell. Syst. Image Process. 2018: 205-212.

[10] Yuan W, Dong S, Adelson E H. Gelsight: High-resolution robot tactile sensors for estimating geometry and force[J]. Sensors, 2017, 17(12): 2762.

[11] Cao G, Zhou Y, Bollegala D, et al. Spatio-temporal attention model for tactile texture recognition[C]//2020 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS). IEEE, 2020: 9896-9902.

[12] Liu Z, Lin Y, Cao Y, et al. Swin transformer: Hierarchical vision transformer using shifted windows[C]//Proceedings of the IEEE/CVF international conference on computer vision. 2021: 10012-10022.

[13] Garcia-Garcia A, Zapata-Impata B S, Orts-Escolano S, et al. Tactilegcn: A graph convolutional network for predicting grasp stability with tactile sensors[C]//2019 International Joint Conference on Neural Networks (IJCNN). IEEE, 2019: 1-8.

[14] Li J, Dong S, Adelson E. Slip detection with combined tactile and visual information[C]//2018 IEEE International Conference on Robotics and Automation (ICRA). IEEE, 2018: 7772-7777.

[15] Zhang Y, Kan Z, Tse Y A, et al. Fingervision tactile sensor design and slip detection using convolutional lstm network. arXiv 2018[J]. arXiv preprint arXiv:1810.02653.

[16] Zhao L, Song Y, Zhang C, et al. T-gcn: A temporal graph convolutional network for traffic prediction[J]. IEEE transactions on intelligent transportation systems, 2019, 21(9): 3848-3858.

[17] Taunyazov T, Chua Y, Gao R, et al. Fast texture classification using tactile neural coding and spiking neural network[C]//2020 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS). IEEE, 2020: 9890-9895.

[18] Gupta A K, Nakagawa-Silva A, Lepora N F, et al. Spatio-temporal encoding improves neuromorphic tactile texture classification[J]. IEEE Sensors Journal, 2021, 21(17): 19038-19046.

[19] A. Le Guennec, S. Malinowski, and R. Tavenard, “Data augmentation for time series classification using convolutional neural networks,” in Proc. AALTD, 2016, pp. 1–9.

[20] T. T. Um et al., “Data augmentation of wearable sensor data for parkin-son’s disease monitoring using convolutional neural networks,” in Proc.19th ACM Int. Conf. Multimodal Interact., Nov. 2017, pp. 216–220.

[21] Du J, Zhang S, Wu G, et al. Topology adaptive graph convolutional networks[J]. arXiv preprint arXiv:1710.10370, 2017.

[22] Davies M, Srinivasa N, Lin T H, et al. Loihi: A neuromorphic manycore processor with on-chip learning[J]. Ieee Micro, 2018, 38(1): 82-99.

[23] Shrestha S B, Orchard G. Slayer: Spike layer error reassignment in time[J]. Advances in neural information processing systems, 2018, 31.

[24] Vaswani A, Shazeer N, Parmar N, et al. Attention is all you need[J]. Advances in neural information processing systems, 2017, 30.

[25] Z. Cui, W. Chen, and Y. Chen, “Multi-scale convolutional neural networks for time series classification,” 2016, arXiv:1603.06995.