Proceedings of The 2021 IEEE International
Conference on Real-time Computing and Robotics
July 15-19, 2021, Xining, China

Tactile Grasp Stability Classification Based on Graph Convolutional Networks

Tingting Mi[1,2], Dashun Que[1], Senlin Fang[2], Zhenning Zhou[2], Chaoxiang Ye[2],
Chengliang Liu[2], Zhengkun Yi[2,4,*], Xinyu Wu[3,5]
[1]School of Information Engineering, Wuhan University of Technology, Wuhan 430070, China
[2]CAS Key Laboratory of Human-Machine Intelligence-Synergy Systems, Shenzhen Institute of Advanced Technology, Shenzhen 518055, China
[3]Guangdong Provincial Key Lab of Robotics and Intelligent System, Shenzhen Institute of Advanced Technology, Chinese Academy of Sciences, Shenzhen 518055, China
[4]Shenzhen Key Laboratory of Smart Sensing and Intelligent Systems, Shenzhen Institute of Advanced Technology, Chinese Academy of Sciences, Shenzhen 518055, China
[5]SIAT Branch, Shenzhen Institute of Artificial Intelligence and Robotics for Society, Shenzhen 518055
*Corresponding author: Zhengkun Yi, zk.yi@siat.ac.cn

2021 IEEE International Conference on Real-time Computing and Robotics (RCAR) | 978-1-6654-3678-6/21/$31.00 ©2021 IEEE | DOI: 10.1109/RCAR52367.2021.9517085

***Abstract*—One of the challenges for robots to grasp unknown objects is to predict whether objects will fall at the beginning of grasping. Evaluating robotic grasp state accurately and efficiently is a significant step to address this issue. In this paper, based on the different fusion approaches of multi-sensor tactile signals, we propose two novel methods based on Graph Convolution Network (GCN) for robotic stability classification. Specifically, we propose two deep learning methods including GCN based on data-level fusion (GCN-DF) and GCN based on feature-level fusion (GCN-FF). We explore the optimal parameters for transforming sensor signals into a graph structure. Furthermore, we verify the effectiveness of the proposed methods on the BioTac Grasp Stability (BiGS) dataset. The experimental results prove that the proposed approaches achieve higher classification accuracy than Support Vector Machine (SVM) and Long Short-Term Memory (LSTM).**

***Index Terms*—Robotic Grasping, Stability Classification, Tactile Sensors, Graph Convolutional Network.**

I. Introduction

Due to the increasing demand of industrial and service robots for various dexterous grasping and manipulation, the grasping task of robots has attracted more and more attention. Human beings can easily adjust the force according to their own feelings when lift an object, which is a challenge for robots. For this problem, we can help robots avoid a probable grasp failure and get a chance to re-grasp the object stably by a reliable prediction of grasp states. Grasp results are usually divided into two states: stable and slippery, in which the former indicates that the object is grasped firmly and the latter represents that the object could fall off the gripper. Grasp stability analysis based on multi-fingered dexterous robotic hand is a research hotspot in the field of intelligent robots. Yamada et al. [1] studied the grasp stability of two objects in three dimensions to increase the flexibility of multi-fingered robot hands. Natarajan et al. [2] designed a three-fingered soft hand on the basis of protecting objects and decreasing costs, and carried out an experiment on gripping force and stability analysis.

The information provided by robots to human beings is mainly visual feedback and tactile feedback in the process of manipulation [3]. We can obtain more accurate properties from tactile perception, such as roughness, contact strength, temperature, pressure, and other invisible details [4]. Recently, the human-inspired biomimmetic tactile sensor (BioTac) capable of simulating rich perceptions of human fingertips has been developed. Chebotar et al. [5] used the Barrett hand equipped with three BioTac sensors to collect tactile data and provide a publicly accessible grasp stability dataset. This is the BiGS data set employed in this paper.

Tactile sensing during the interaction with target objects is widely used in predicting grasp stability. Dang et al. [6] made use of tactile feedback to predict the stability of a robotic grasp without any visual or geometric information about the grasped object. Bekiroglu et al. [7] investigated the influence of different sensory streams on grasp stability, and finally proved that the grasp results can be inferred by learning the tactile measurements from fingertips and joint configuration of the hand. Cockbum et al. [8] was committed to exploring the common denominators behind successful and failed grasps, and they evaluated grasp quality by exploiting only high-level features from two tactile sensors. In a word, tactile sensing plays a prominent role in the scientific research of robot perception.

For grasp stability classification, the existing methods explore various machine learning algorithms to extract features from tactile sensors. A typical approach is to send tactile representations to a Support Vector Machine (SVM) classifier [9]. In recent years, a variety of deep learning techniques suitable for processing large-scale data are being applied to automatic classification tasks. Begalinova et al. [10] adopted Long Short-Term Memory (LSTM) networks to classify the time-series pressure readings from low-cost tactile sensors as either slip or non-slip events. Futhermore, Zhang et al. [11] proposed a convolution Long Short-Term Memory (ConvL-

STM) network for slip recognition, which made a remarkable effect on the improvement of grasp stability. The previous works suffer from a common drawback. To be specific, the related works treat tactile readings as regular array or matrix signals. However, the neglected inherent correlations between each tactile sensing point and its adjacent points will probably exert an influence on stability prediction.

We propose two Graph Convolution Network (GCN) models based on data-level fusion and feature-level fusion respectively, to carry out the grasp stability classification. The premise of this work is to express the raw tactile data as a graph structure. And the main challenge comes from implementing a convolution operation on the tactile graph. Some contributions made in this work can be summarized as follows.

1) We transform the electrode values of BioTac sensors into graph representations by using k-Nearest Neighbors (kNN) graph algorithm. Moreover, the optimization of k values further improve the classification performance.
2) We complete the convolution operation on the constructed graph, which is capable of learning the real spatial distribution information of sensing nodes.
3) We present two fusion strategies including data-level fusion and feature-level fusion to integrate the features of multiple sensors. Compared with the single sensor model, the models with fusion methods perform better for stability recognition.
4) We perform extensive comparison experiments on a publicly available grasp stability dataset to prove the validity of the proposed GCN methods.

We organize the rest of this paper in the following order. Section II gives the detailed theories and steps of the proposed methods. In section III, we provide an introduction to dataset and the experimental results. At last, section IV presents the conclusion and the future work.

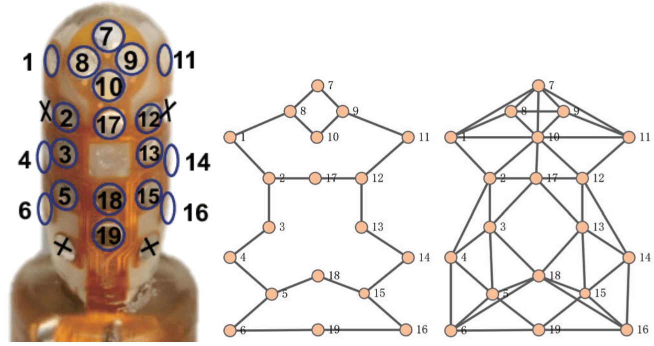
II. METHODS

In this section, we first introduce the construction approach of tactile graph in Section II-A and provide the principle of graph convolution operation in Section II-B. After that, the architectures of the proposed networks are described in detail in Section II-C.

A. Tactile Graph Construction

The tactile data to be processed by our proposed Graph Convolutional Network (GCN) is BioTac electrode values. For the common CNN model, the input data is represented as a regular grid structure. In contrast, sensor readings need to be expressed as graph representation [12] due to the particularity of GCN in our model. Therefore, we utilize the position distribution of 19 electrodes on the BioTac sensor to construct tactile graph, which is an essential step in our work. The 3D coordinates $(\mathbf{x}_i, \mathbf{y}_i, \mathbf{z}_i)$ of the 19 sensing electrodes [13] are provided in Table I.

Specifically, let the haptic graph be represented as a tuple $\mathbf{G} = (\mathrm{V}, \mathrm{E})$, where $\mathbf{V} = \{\mathrm{v}_1, \mathrm{v}_2, \cdots, \mathrm{v}_N\}$ refers to a set of $\mathbf{N}$ electrode nodes, and $\mathbf{E}$ means a set of undirected edges



(a) BioTac sensor (b) k=2 (c) k=4

Fig. 1: The electrode values of BioTac sensors are transformed into graph representations by using kNN algorithm. (a) The BioTac sensor with 19 sensing electrodes and 4 excitation electrodes. (b) Graph generated by kNN (k = 2). (c) Graph constructed by kNN (k = 4).

connecting nodes. The feature vector, namely the electrode values detected by the sensor during the grasping process, contained in node $\mathbf{v}_i$ is $\mathbf{f}_i = (\mathrm{f}_{i1}, \mathrm{f}_{i2}, \cdots, \mathrm{f}_{iC})$, where $\mathbf{C}$ is the number of features.

Given the specific coordinates of each node, k-Nearest Neighbors (kNN) graph algorithm [14] can easily construct tactile graph. For a node $\mathbf{v}_i$, we calculate the Euclidean distance between it and other nodes to find out the k nearest neighbors, and then connect $\mathbf{v}_i$ with these neighbors to generate k directed edges. In practice, we treat directed edges in kNN graphs as undirected edges. Figure 1 shows the tactile graphs constructed by adopting the k-Nearest Neighbors (kNN) method.

B. Convolutions on Graphs

Convolutional Neural Network (CNN) shows its powerful feature extraction and integration ability in processing data with regular grid structure, such as 2D image and 1D speech. Corresponding to CNN, GCN explores how to apply convolution operation to topological graph. Convolution on graphs

TABLE I: Electrode position in 3D coordinates

Coordinates	$\mathbf{x}(mm)$	$\mathbf{y}(mm)$	$\mathbf{z}(mm)$
$\mathbf{E}_1$	0.993	-4.855	-1.116
$\mathbf{E}_2$	-2.700	-3.513	-3.670
$\mathbf{E}_3$	-6.200	-3.513	-3.670
$\mathbf{E}_4$	-8.000	-4.956	-1.116
$\mathbf{E}_5$	-10.500	-3.513	-3.670
$\mathbf{E}_6$	-13.400	-4.956	-1.116
$\mathbf{E}_7$	4.763	0.000	-2.330
$\mathbf{E}_8$	3.031	-1.950	-3.330
$\mathbf{E}_9$	3.031	1.950	-3.330
$\mathbf{E}_{10}$	1.299	0.000	-4.330
$\mathbf{E}_{11}$	0.993	4.855	-1.116
$\mathbf{E}_{12}$	-2.700	3.513	-3.670
$\mathbf{E}_{13}$	-6.200	3.513	-3.670
$\mathbf{E}_{14}$	-8.000	4.956	-1.116
$\mathbf{E}_{15}$	-10.500	3.513	-3.670
$\mathbf{E}_{16}$	-13.400	4.956	-1.116
$\mathbf{E}_{17}$	-2.800	0.000	-5.080
$\mathbf{E}_{18}$	-9.800	0.000	-5.080
$\mathbf{E}_{19}$	-13.600	0.000	-5.080

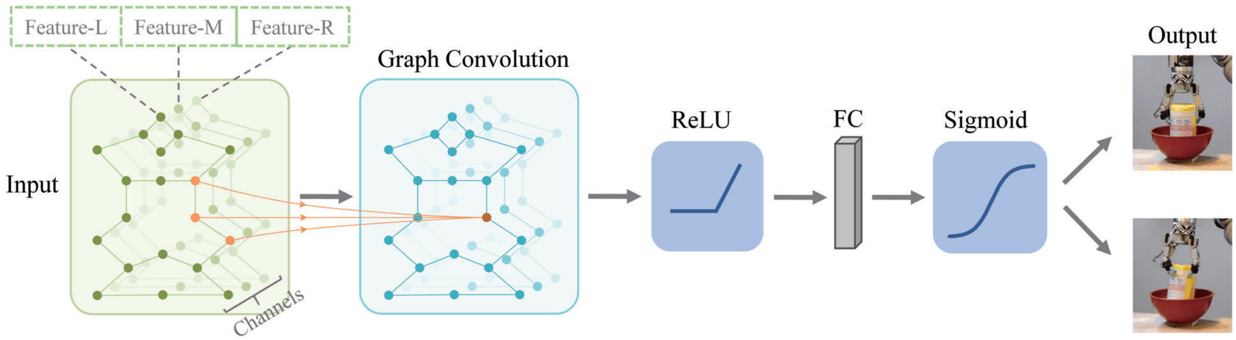


Fig. 2: GCN-DF model for robotic stability classification. GCN-DF is a Graph Convolution Network based on data-level fusion. In the input graph, each node contains the electrode values of multiple sensors, where Feature-L, Feature-M and Feature-R respectively denote the readings of tactile sensors equipped in the left, middle, and right finger. Output is divided into two grasp results: a stable state (upper right) and a sliding state (lower right).

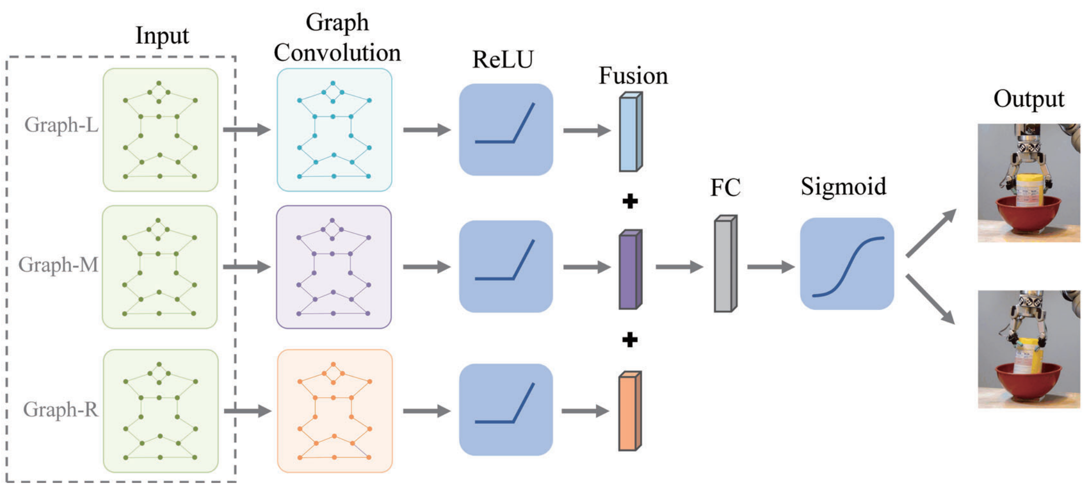


Fig. 3: GCN-FF architecture. GCN-FF utilizes the feature-level fusion means to integrate the information of multiple sensors. Data collected by 3 mechanical fingers are converted into Graph-L, Graph-M and Graph-R.

includes spectral domain graph convolution and spatial domain graph convolution [15]. The former has a solid theoretical basis for processing data in spectral domain, while the latter directly defines convolution operation in space. In this paper, we choose the spatial approach considering its intuitive definition and high flexibility.

We use the adjacency matrix $\mathbf{A} \in \mathbb{R}^{N\times N}$ to describe the association between nodes, and define a degree matrix $\mathbf{D}_{ii} = \sum_j \mathbf{A}_{ij}$ to summarize the number of edges connected by each node. The adjacency matrix $\mathbf{A}$ is a diagonally symmetric matrix with its elements denoted as:

$$\mathbf{A}_{ij} = \begin{cases} 1 & \text{if } (\mathrm{v}_i, \mathrm{v}_j) \subseteq \mathrm{E} \\ 0 & \text{else.} \end{cases} \tag{1}$$

The aggregation and update of node features are implemented by constructing a neural network model $\mathbf{H}^{(l+1)} = \mathbf{f}\left(\mathbf{H}^{(l)}, \mathbf{A}\right)$. $\mathbf{H}^{(l)}$ represents the feature matrix of the l^{th} layer and $\mathbf{H}^{(0)}$ is the initial feature map $\mathbf{X} \in \mathbb{R}^{N\times C}$ composed of Biotac electrode values. The propagation rule that we used to build our GCN model can be expressed with the following formula [13]:

$$\mathbf{H}^{(l+1)} = \sigma\left(\widehat{\mathbf{D}}^{-\frac{1}{2}}\widehat{\mathbf{A}}\widehat{\mathbf{D}}^{-\frac{1}{2}}\mathbf{H}^{(l)}\mathbf{W}^{(l)}\right) \tag{2}$$

where $\widehat{\mathbf{A}} = \mathbf{A} + \mathbf{I}$ is a new adjacency matrix with added an identity matrix $\mathbf{I}$ to realize self-connections, $\widehat{\mathbf{D}}_{ii} = \sum_j \widehat{\mathbf{A}}_{ij}$ and $\mathbf{W}^{(l)}$ denotes a learnable weight matrix at the l^{th} layer. In addition, we use ReLU as the activation function $\sigma(\cdot)$ in our network.

C. Model Architecture

We validate the effectiveness of our classification methods on the dataset BiGS, which was collected using the 3-finger Barrett WAM^{TM} Arm manipulator and each finger equipped with a BioTac sensor. In order to integrate the tactile information of three sensors, we explored two data fusion methods, including data-level fusion and feature-level fusion [16], [17]. On this basis, We present a GCN model based on data-level

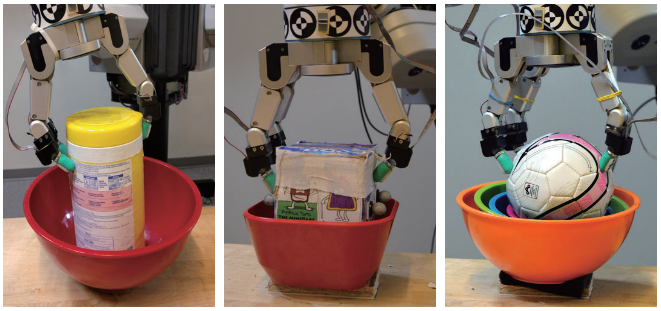


Fig. 4: Experimental setup for collecting the grasp stability dataset. The objects used in the experiment are a cylindrically-shaped box of wipes (left), a cubically-shaped box of candy (middle) and a ball (right).

fusion (GCN-DF) and a GCN model based on feature-level fusion (GCN-FF).

The GCN-DF splices the raw electrode signals of 3 sensors into a feature vector (Feature-L, Feature-M, Feature-R) at their corresponding nodes. Feature-L, Feature-M, Feature-R are feature vectors associted with the left, middle, and right finger, respectively. As show in Figure 2, each node of the input haptic graph contains data of 3 mechanical fingers. GCN-FF sends the tactile graphs Graph-L, Graph-M and Graph-R of 3 fingers to GCN respectively, and then concatenates the extracted feature vectors for classification, which belongs to feature-level fusion method. Figure 3 shows the overall structure of the proposed GCN-FF model.

On the whole, these two stability classification methods have similar structures. They both adopt one-layer GCN to update the node features and use ReLU as the activation function. At the end of the model, we use a fully connected (FC) layer with sigmoid to carry out a binary classification. The output is divided into two grasping states: a successful grasp and an unsuccessful grasp.

III. Experiments

In this section, we detail the experimental device and data collection process of BiGS dataset. The proposed models are compared with two previous techniques to verify the superior performance. Finally, we calculate the influence of connectivity on GCN and utilize the recognition accuracy and F1 score to evaluate the performance of all models.

A. BiGS Dataset

In this paper, the effectiveness of our recognition methods is demonstrated on the BioTac Grasp Stability (BiGS) dataset which collected by Chebotar et al. from the University of Southern California. The Barrett WAM^{TM} Arm manipulator and three BioTac sensors installed on the Barrett hand were employed to record multimodal data in the process of grasping objects. In addition to the BioTac electrode values involved in our experiments, BiGS also contains some other modalities, such as BioTac pressure sensor values, robot's joint angles, finger joint angles and object pose obtained from the VICON system. Sampling rate of the raw BioTac data is 100Hz, except AC pressure signal (PAC) which is sampled at 2200 Hz, and the robot data is sampled at 300Hz.

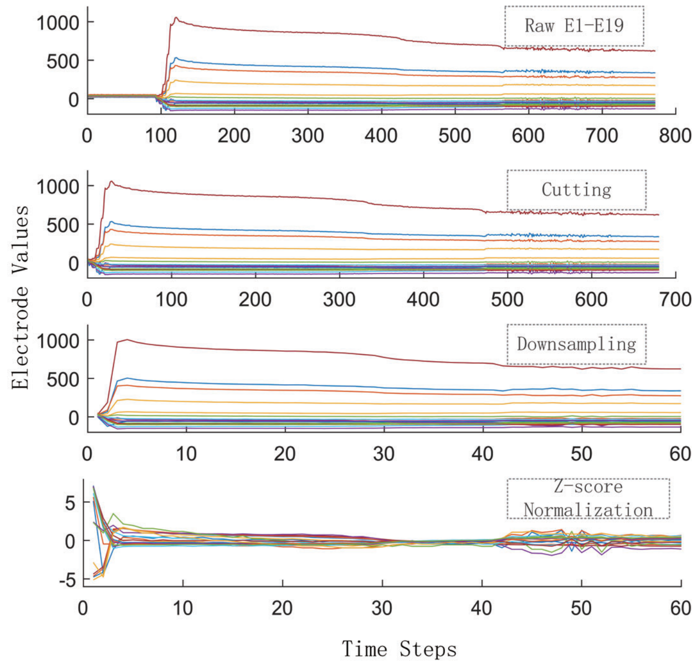


Fig. 5: Data preprocessing. Each sample contains 19 BioTac electrode signals. The preprocessing of tactile signals can be divided into three steps: cutting, downsampling and normalization.

Chebotar et al. manipulated a haptically-enabled robot to grasp three common objects with different shapes: a cylindrically-shaped box of wipes, a cubically-shaped box of candy and a ball. They used a VICON motion-tracking system to track the target object and calculate its initial position in advance. The specific data acquisition process is as follows. The robotic arm gradually approaches the object and performs a random top grasp with a force grip controller. In order to ensure the stability of grasping, the robot will conduct a range of extensive shaking movements in all directions after lifting the object up. If the object remains in the hand after shaking, it is marked as a successful grasp. In addition, the experimenters fixed a bowl on the table to catch the object if it falls out of the gripper. The experimental setup is displayed in Figure 4.

The BiGS dataset provides a total of 2000 grasps. There are 54% successes and 46% failures in 1000 grasps of the cylinder object. The box object was tested 500 times, 69% of them failed and 31% succeeded. For the ball, the robot gathered 500 grasps, out of which 47.4% resulted in failures and 52.6% succeeded.

B. Implementation Details

The time sequences provided by BiGS dataset begin with the moment the fingers start closing around the object and end 2 seconds after starting picking up the object. We discard the frames from the beginning of the raw time series to the point when the fingers touch the object, because the information contained in this period contributes less to our classification

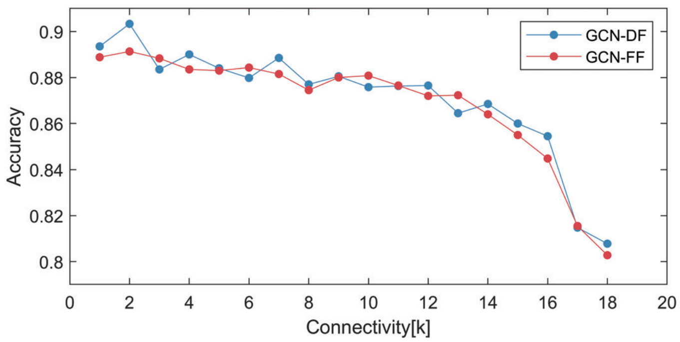


Fig. 6: Validation accuracy curves of GCN models according to different connectivity of the k-NN graph.

task. In order to further improve the training efficiency, we down-sample the input to 60 frames, and normalize it by using Z-score normalization. Detailed data preprocessing steps are given in Figure 5.

We mix and randomly shuffle all samples of the cylinder, the box and the ball, and divide the 2000 grasps into training and testing sets with the ratio of 8:2. In this paper, all deep learning models are trained for 100 epochs, with a batch size of 32 samples. The Adam optimizer with a learning rate of 0.001 is used to minimize the cross-entropy loss function.

C. Compared Methods

We compare our proposed models with the existing state-of-the-art methods: Support Vector Machine (SVM) and Long Short-Term Memory (LSTM). All compared methods are verified on the same dataset. A brief introduction of these two techniques is as follows.

1) SVM: SVM is a popular supervised machine learning model, which is widely used to solve linear and non-linear classification problems [18]–[20]. We extract 12 features [21] from tactile time series to form a fixed-length feature representation as the input of SVM. The hand-crafted features consist of maximum, minimum, peak, mean, variance, standard deviation, absolute mean, root mean square, crest factor, form factor, pulse factor, and margin factor. Since each sample contains readings of 19 electrodes, the length of the generated feature vector is 228.
2) LSTM: By taking the hidden layer as a memory unit [22], LSTM can effectively learn the correlation within time series in both short and long term. In our comparison experiments, one-layer LSTM network whose hidden layer includes 128 memory cells is selected for stability classification. The output layer is sigmoid, and the results correspond to two grasp states.

We evaluate the performance of single sensor and multi-sensor fusion with SVM, LSTM and GCN respectively. All results are described in the next section.

D. Experiment Results

The complexity of k-NN graph is affected by k, and finding an optimal k value is beneficial to improve the performance of GCN. Referring to [13], we regard k as graph connectivity. Obviously, the larger k means the stronger connectivity. Since the tactile graph contains 19 nodes and each point is attached to its k nearest neighbors, we investigate the classification capability of the k-NN strategy with $k \in \{1, 2, \cdots, 18\}$ on GCN-DF and GCN-FF.

The accuracy curves are shown in Figure 6, from which we see that $k = 2$ has the best performance on both networks. In general, larger k values achieve worse performance in terms of accuracy. The reason for this trend is that excessive k values will weaken or even eliminate the distance variance between nodes.

To prove the effectiveness of our methods, we compare the proposed methods with SVM and LSTM. The results of the proposed approaches are obtained with $k = 2$. Table II illustrates all experimental results. Each value is the average of 10 repeated experiments on the test set. The suffixes ”-L”, ”-M” and ”-R” indicate that the input data comes from the sensors assembled on the left, middle and right fingers respectively. ”-F” denotes that data fusion is applied. We use recognition accuracy (Acc) and F1 score (F1) as metrics to measure the performance of all models. F1 score is the harmonic mean of precision (Prec) and recall (Rec). These metrics are formulated as:

$$\mathrm{Acc} = \frac{\mathrm{TN} + \mathrm{TP}}{\mathrm{TN} + \mathrm{TP} + \mathrm{FN} + \mathrm{FP}} \tag{3}$$

$$\mathrm{Prec} = \frac{\mathrm{TP}}{\mathrm{TP} + \mathrm{FP}} \tag{4}$$

$$\mathrm{Rec} = \frac{\mathrm{TP}}{\mathrm{TP} + \mathrm{FN}} \tag{5}$$

$$\mathrm{F1} = 2\frac{\mathrm{Prec} \cdot \mathrm{Rec}}{\mathrm{Prec} + \mathrm{Rec}} \tag{6}$$

where TP is the true positive rate, FP is the false positive rate, and FN is the false negative rate.

It is observed that GCN models achieve higher accuracy and F1 score than previous state-of-the-art models in the case of single sensor and multi-sensor. The reason is that GCN can learn node features and structure information end to end at the same time, while SVM and LSTM ignore the real spatial distribution of electrodes. Additionally, deep learning methods LSTM and GCN perform better than SVM with hand-crafted

TABLE II: The comparison results of different classification methods

Methods	**Accuracy (%)**	**F1 score (%)**
SVM-L	79.50	78.42
SVM-M	80.75	79.79
SVM-R	79.75	79.90
SVM-F	85.50	84.97
LSTM-L	82.35	81.87
LSTM-M	84.68	84.36
LSTM-R	82.23	81.75
LSTM-F	86.20	85.62
GCN-L	83.08	82.14
GCN-M	87.00	86.86
GCN-R	85.10	85.41
GCN-DF	**90.33**	**90.08**
GCN-FF	**89.13**	**88.79**

features. This is probably due to the strong learning ability of deep learning methods for large-scale data. In contrast, using fusion strategy can always further increase the classification performance because multiple sensors contain more information than a single sensor. Especially, the proposed fusion method GCN-DF obtains the highest accuracy of 90.33%, followed by GCN-FF with an accuracy of 89.13%.

IV. Conclusion

In this paper, we explore two GCN networks based on different fusion strategies for robotic grasp stability classification. We first convert the electrode values of BioTac sensors into graph forms by the k-NN graph algorithm. To select a optimal parameter for graph construction, we carry out extensive experiments by adjusting k within a certain range to find out a highest accuracy. We believe that the data of multiple sensors can provide more effective information for recognition than a single sensor. We confirm this viewpoint by adopting two fusion methods, which perform better than the single sensor model. In addition, we compare the existing state-of-the-art machine learning methods with our methods on the same tactile dataset. The experimental results demonstrate that the proposed networks achieve higher recognition accuracy. In the future work, we plan to combine other modalities, such as force torque and finger joint angles, to improve the classification performance.

Acknowledgment

The work described in this paper is partially supported by the National Natural Science Foundation of China (Grant No. 62003329), partially supported by the NSFC-Shenzhen Robotics Research Center Project (Grant No. U2013207), and partially supported by the National Natural Science Foundation of Guangdong Province of China (Grant No. 2021A1515011316). The authors are grateful to the financial support from the Science and Technology Innovation Commission of Shenzhen (No. ZDSYS20190902093209795).

References

[1] T. Yamada, T. Taki, M. Yamada, and H. Yamamoto, "Grasp stability analysis of two objects by considering contact surface geometry in 3d," in *2011 IEEE International Conference on Robotics and Biomimetics*, pp. 1108–1115, IEEE, 2011.

[2] E. Natarajan, M. I. Solihin, and J. H. Chong, "Grasp stability analysis of an isotropic direct driven three-finger soft robot hand," *International Journal on Advanced Science, Engineering and Information Technology*, vol. 7, no. 5, pp. 1627–1631, 2017.

[3] D. Guo, F. Sun, B. Fang, C. Yang, and N. Xi, "Robotic grasping using visual and tactile sensing," *Information Sciences*, vol. 417, pp. 274–286, 2017.

[4] S. Cui, R. Wang, J. Wei, F. Li, and S. Wang, "Grasp state assessment of deformable objects using visual-tactile fusion perception," in *2020 IEEE International Conference on Robotics and Automation (ICRA)*, pp. 538–544, IEEE, 2020.

[5] Y. Chebotar, K. Hausman, Z. Su, A. Molchanov, O. Kroemer, G. Sukhatme, and S. Schaal, "Bigs: Biotac grasp stability dataset," in *ICRA 2016 Workshop on Grasping and Manipulation Datasets*, 2016.

[6] H. Dang and P. K. Allen, "Learning grasp stability," in *2012 IEEE International Conference on Robotics and Automation*, pp. 2392–2397, IEEE, 2012.

[7] Y. Bekiroglu, J. Laaksonen, J. A. Jorgensen, V. Kyrki, and D. Kragic, "Assessing grasp stability based on learning and haptic data," *IEEE Transactions on Robotics*, vol. 27, no. 3, pp. 616–629, 2011.

[8] D. Cockbum, J.-P. Roberge, A. Maslyczyk, V. Duchaine, *et al.*, "Grasp stability assessment through unsupervised feature learning of tactile images," in *2017 IEEE International Conference on Robotics and Automation (ICRA)*, pp. 2238–2244, IEEE, 2017.

[9] J. W. James, N. Pestell, and N. F. Lepora, "Slip detection with a biomimetic tactile sensor," *IEEE Robotics and Automation Letters*, vol. 3, no. 4, pp. 3340–3346, 2018.

[10] A. Begalinova, R. D. King, B. Lennox, and R. Batista-Navarro, "Self-supervised learning of object slippage: An lstm model trained on low-cost tactile sensors," in *2020 Fourth IEEE International Conference on Robotic Computing (IRC)*, pp. 191–196, IEEE, 2020.

[11] Y. Zhang, Z. Kan, Y. A. Tse, Y. Yang, and M. Y. Wang, "Fingervision tactile sensor design and slip detection using convolutional lstm network," *arXiv preprint arXiv:1810.02653*, 2018.

[12] Z. Yi, W. Shang, T. Xu, and X. Wu, "Neighborhood Preserving and Weighted Subspace Learning Method for Drift Compensation in Gas Sensor," *IEEE Transactions on Systems, Man, and Cybernetics: Systems (Early Access)*, pp. 1–12, 2021.

[13] A. Garcia-Garcia, B. S. Zapata-Impata, S. Orts-Escolano, P. Gil, and J. Garcia-Rodriguez, "Tactilegcn: A graph convolutional network for predicting grasp stability with tactile sensors," in *2019 International Joint Conference on Neural Networks (IJCNN)*, pp. 1–8, IEEE, 2019.

[14] F. Gu, W. Sng, T. Taunyazov, and H. Soh, "Tactilesgnet: A spiking graph neural network for event-based tactile object recognition," *arXiv preprint arXiv:2008.08046*, 2020.

[15] B. Yu, H. Yin, and Z. Zhu, "Spatio-temporal graph convolutional networks: A deep learning framework for traffic forecasting," *arXiv preprint arXiv:1709.04875*, 2017.

[16] R. Gravina, P. Alinia, H. Ghasemzadeh, and G. Fortino, "Multi-sensor fusion in body sensor networks: State-of-the-art and research challenges," *Information Fusion*, vol. 35, pp. 68–80, 2017.

[17] Q. Liang, Z. Yi, Q. Hu, and Y. Zhang, "Low-cost sensor fusion technique for surface roughness discrimination with optical and piezoelectric sensors," *IEEE Sensors Journal*, vol. 17, no. 23, pp. 7954–7960, 2017.

[18] Z. Yi, T. Xu, S. Guo, W. Shang, and X. Wu, "Tactile Surface Roughness Categorization With Multineuron Spike Train Distance ," *IEEE Transactions on Automation Science and Engineering (Early Access)*, pp. 1–11, 2020.

[19] Z. Yi, W. Shang, T. Xu, S. Guo, and X. Wu, "Local Discriminant Subspace Learning for Gas Sensor Drift Problem," *IEEE Transactions on Systems, Man, and Cybernetics: Systems (Early Access)*, pp. 1–13, 2020.

[20] Z. Yi, T. Xu, W. Shang, and X. Wu, "Touch Modality Identification with Tensorial Tactile Signals: A Kernel-based Approach ," *IEEE Transactions on Automation Science and Engineering (Early Access)*, pp. 1–10, 2021.

[21] H. Wu, X. Liu, S. Fang, Z. Yi, and X. Wu, "Leveraging multi-label correlation for tactile adjective recognition," in *2020 3rd International Conference on Robotics, Control and Automation Engineering (RCAE)*, pp. 122–126, IEEE, 2020.

[22] F. Pastor, J. García-González, J. M. Gandarias, D. Medina, P. Closas, A. J. García-Cerezo, and J. M. Gómez-de Gabriel, "Bayesian and neural inference on lstm-based object recognition from tactile and kinesthetic information," *IEEE Robotics and Automation Letters*, vol. 6, no. 1, pp. 231–238, 2020.