Response to Review #258645

The motivation and contribution of the work are well presented in the introduction section. Authors have used three datasets to support their results with other counterparts. The overall presentation of the paper is good and complete.

I have following concerns related to this work:

Thank you very much for your affirmation and constructive comment. The point-by-point responses to the comments are provided as follows.

1) Authors mentioned about constructing different local graphs for different training batches, so only one concern is how the change in the batch size affects the performance of the algorithm.

RESPONSE:

To address your concern, we have implemented some experiments on the classification results with different batch sizes *s*, which are shown as follows:

s	Accuracy-mean(%)			F1-Score-mean(%)		
	MNIST	Fashion	CIFAR10	MNIST	Fashion	CIFAR10
50	99.09	96.70	74.48	98.88	96.28	73.29
100	99.48	95.71	73.62	99.27	95.19	72.31
150	99.24	94.86	73.45	99.03	94.51	72.16

From the classification results of GDGCNN with different batch sizes, we can see that different batch sizes s should be chosen on different datasets to obtain the best classification results. In addition, GDGCNN is not sensitive to s and has strong robustness on these three datasets.

2) Does the proposed approach be able to handle imbalance classification problems? The class distribution for different data sets is not provided in the paper. What would be the impact of class imbalance problem?

RESPONSE:

Thank you very much for your valuable comment. We did not study this problem, but we think that GDGCNN considers feature learning in each batch, which would be better than GCNN that only considers global features. We think the issue you raised is a good direction, and we would like to study it in our future algorithms.

3) More data sets and more comparing methods should be introduced in the experimental result section.

RESPONSE:

To address your concern, we have performed a classification comparison experiment of 1stChebNet [Thomas and Welling, 2017], StoGCN [Chen et al., 2018] and GDGCNN on three common datasets [Thomas and Welling, 2017; Chen et al., 2018]. The ACC results are shown as follows:

	Citeseer	Cora	Pubmed
1stChebnet	70.3%	81.5%	**79.0%**
StoGCN	**70.9%**	82.0%	**79.0%**
GDGCNN	70.7%	**82.3%**	**79.0%**

From the above ACC results, we can see that GDGCNN has better classification results than 1stChebnet [Thomas and Welling, 2017] on Citeseer and Cora. In addition, the ACC of GDGCNN is higher than StoGCN on Cora. Therefore, it shows that GDGCNN has some advantages some over state-of-the-art methods on these three common datasets.

References:

[Thomas and Welling, 2017] N Kipf Thomas and Max Welling. Semi-supervised classification with graph convolutional networks. in Proceedings of the International Conference on Learning Representations, 2017.

[Chen et al., 2018] Jianfei Chen, Jun Zhu, and Le Song. Stochastic training of graph convolutional networks with variance reduction. in Proceedings of the International Conference on Machine Learning, 2018, pp. 941– 949.

4) The samples in training batches are randomly chosen, not necessarily the samples are local neighbors. A study on how the algorithm handles this situation may enrich the work.

RESPONSE:

We also agree with you that the study you pointed out is a good direction, and we would like to study it in the future.

5) Parameter analysis is missing in the result section.

RESPONSE:

Considering the body length requires less than 6 pages, we did not perform the parameter analysis, and the classification results with the main parameters the neighbor number k and batch size s as follows:

k	Accuracy-mean(%)			F1-Score-mean(%)		
	MNIST	Fashion	CIFAR10	MNIST	Fashion	CIFAR10
2	99.36	96.99	30.26	99.13	96.58	30.11
5	99.56	96.85	63.74	99.35	96.48	62.53
8	99.48	96.70	74.48	99.27	96.28	73.29

From the classification results of GDGCNN with different batch sizes and neighbor numbers, we can see that GDGCNN is not sensitive to k on MNIST and Fashion, and sensitive to k on CIFAR10.

s	Accuracy-mean(%)			F1-Score-mean(%)		
	MNIST	Fashion	CIFAR10	MNIST	Fashion	CIFAR10
50	99.09	96.70	74.48	98.88	96.28	73.29
100	99.48	95.71	73.62	99.27	95.19	72.31
150	99.24	94.86	73.45	99.03	94.51	72.16

From the classification results of GDGCNN, we can see that different batch sizes s should be chosen on different datasets to obtain the best classification results. In addition, GDGCNN is not sensitive to s and has strong robustness on these three datasets.

6) For a suggestion, authors can include some qualitative analysis.

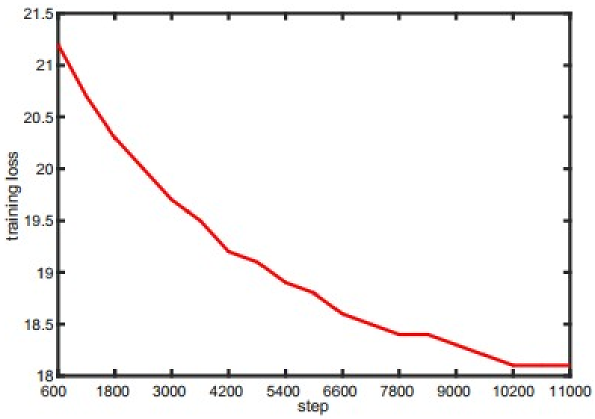
RESPONSE:

To address your concern, we have added some explanations in "IV. C. Parameter Analysis" as follows, which greatly improves the quality of the manuscript.

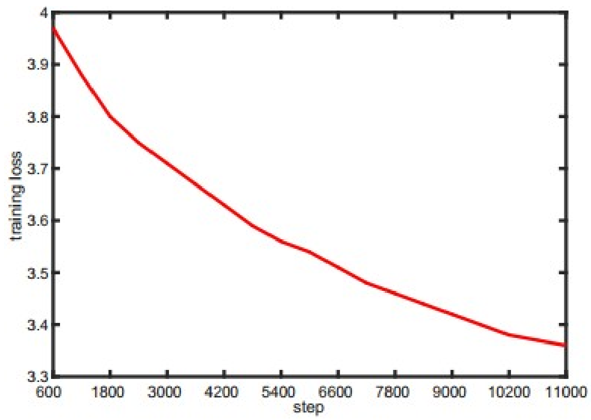
7) Author(s) reported training loss for only MNIST dataset. They should report both training and testing loss for all the three datasets. It will reflect the overall performance of the model..

RESPONSE:

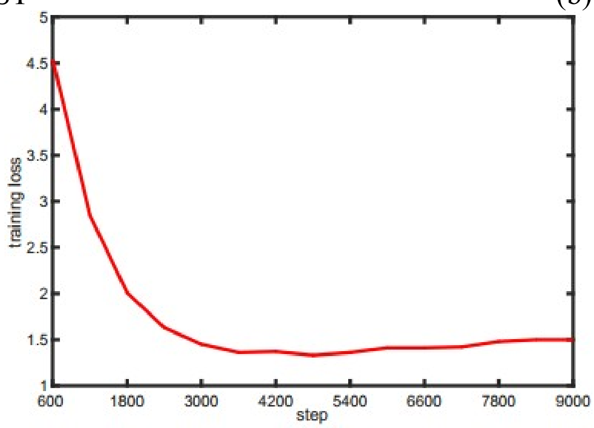
In fact, we have performed convergence experiments on all three datasets, but we only select one image to display in our paper due to the page limit. Here we mainly show the convergence of the proposed GDGCNN on the three datasets (MNIST, Fashion and CIFAR10) as follows:



(a) MNIST



(b) Fashion



(c) CIFAR10

For each image, the abscissa represents the number of iterations and the ordinate represents the training loss. It can be seen that the training loss of GDGCNN has a decreasing trend on all the three datasets, which can achieve convergence. The convergence speed is slow on MNIST and Fashion, which needs more than 10,000 iterations to converge. On CIFAR 10, it can converge only in about 3000 iterations.

Thank you very much for your insightful comments, which have greatly helped us in improving the quality of the paper. Thank you!