GuCNet: A Guided Clustering-based Network for Improved Classification

Ushasi Chaudhuri^a, Syomantak Chaudhuri^b & Subhasis Chaudhuri^b

^aCenter of Studies in Resources Engineering (CSRE) ^bDepartment of Electrical Engineering (EE) Indian Institute of Technology Bombay (IITB)

International Conference on Pattern Recognition, 2021

Introduction to Problem

- Classification problem in vision data.
- Aim: Need to extract relevant features from patterns & project it onto an embedding space
- **Ensure:** Representations of each class of patterns are uniquely distinguishable.
- Semantic classification of challenging and highly-cluttered data is difficult.



・ロト ・ 日 ・ ・ 日 ・ ・ 日 ・ ・ 日

Guided Clustering

Many well-separable datasets are available.

Can we leverage the classifiability of any existing well separable dataset?

- Guide data (X): A well separable data.
- Cluttered data (*Y*): The cluttered dataset, which is to be classified.
- Embed class-wise features of the cluttered data to the distinct clusters of the guide data, to make them more separable.
- Therefore, guided-clustering.



1. Separable Texture-based Guiding



A well-separable data acts as a texture data.

- Extract initial level features from both data using a pre-trained network.
- Feed samples of class-*c* of both X and Y together as the same class label in the unified space.
- Minimize cross-entropy $\mathcal{L}_{CE} = CE(p'_x, p'_y)$.

◆□▶ ◆□▶ ◆三▶ ◆三▶ 三三 のへで

2. Separable Prototype-based Guiding



If a well-separable data of *C*-class unavailable, we can also use a **prototype**-based guided clustering.

- Extract initial level features from cluttered data using a pre-trained network.
- Choose K ($K \ge C$) dimensional vectors (called *prototypes* g).
- Matching loss: (L_{ml} = | p'_x g |) to bring the dataset closer to the prototype vectors.
- Minimize $\mathcal{L}_{ml} = |p'_x g| + CE(p'_x, p'_y).$

- ◆ □ ▶ → 個 ▶ → 目 ▶ → 目 → ○ ○ ○

Results - RSSCN aerial scene dataset

Table: Classification performance of the proposed GuCNet architecture on RSSCN dataset. Here baseline for guide data (MNIST) is 99.80%.

Model	Accuracy(%)		
LLC (CH)	79.94%		
SpLSA (SIFT)	79.37%		
VLAD (SIFT)	79.34%		
RGSIR	81.00%		
AlexNet	88.80%		
CaffeNet	88.60%		
GoogleNet	79.80%		
VGG-M	87.30%		
VGG-VD16	85.60%		
Conv5-MSP5-FV	95.40%		
Baseline	88.39%		
GuCNet (Prototype)	97.36%		
GuCNet (Texture)	99.11%		



Results - LSUN Outdoor scene dataset

Table: Classification performance of GuCNet on LSUN dataset with the same guide data.

Model	Accuracy(%)	
Vanilla GAN	70.50%	
Labeled-samples	77.00%	
ds-cube	83.00%	
Hybrid GAN	83.20%	
Normal BN-Inception+scene n/w	90.40%	
Deeper BN-Inception+scene n/w	90.90%	
SJTU-ReadSense	90.40%	
SIAT MMLAB	91.60%	
Baseline	83.75%	
GuCNet (Prototype)	95.03%	
GuCNet (Texture)	94.86%	





 $\mathcal{Y}\text{: LSUN}$



Results - TU-Berlin sketch dataset

Table: Performance comparison on TU-Berlin dataset for classification accuracy. Here baseline accuracy for guide data is 84.54%.

Model	Accuracy(%)		
AlexNet-SVM	67.10%		
AlexNet-Sketch	68.60%		
Sketch-A-Net SC	72.20%		
Sketch-A-Net-Hybrid	73.10%		
ResNet18-Hybrid	73.80%		
Humans	73.10%		
Sketch-A-Net-Hybrid	orid 77.00%		
Sketch-A-Net	77.00%		
Alexnet-FC-GRU	79.95%		
Zhang et. al.	82.95%		
Baseline	69.90%		
GuCNet (Prototype)	86.63%		
GuCNet (Texture)	89.26%		

\mathcal{X} : TU-Berlin Images





 $\mathcal{Y}: \ \mathsf{TU}\text{-}\mathsf{Berlin} \ \mathsf{Sketches}$





◆□ ▶ ◆母 ▶ ◆臣 ▶ ◆臣 ▶ ○臣 ○のへで

Some Interesting Ablation Study:

Table: Effect of different types of **co-binning** of texture classes from guide set.

Dataset (TU-Berlin)	Accuracy(%)	
GuCNet (Texture):		
Same class binning	89.26%	
Dissimilar class binning	90.05%	

Table: Effect of separability of prototypes in terms of Hamming distance (H) on GuCNet performance.

Dataset	Separation of prototypes				
	w2vec	H = 2	$\frac{H_{\text{max}}}{2}$	H _{max}	
GuCNet (Prototype):					
RSCCN dataset	96.20%	96.02%	96.27%	97.36%	
LSUN dataset	92.71%	94.60%	94.92%	95.03%	

Conclusion

- We propose a very simple yet novel guided clustering to get high performance in classification.
- Leverage the ease of separability of a guide dataset to improve the separability of a cluttered dataset.
- Pushes the embeddings of the data instances far apart in the semantic feature space while making the embedding space further discriminative.
- Established its efficacy on three challenging datasets and outperformed the state-of-the-art performance.
- **Future work:** Can try to match the distributions of the two datasets and study the performance.

Thank You

◆□▶ ▲□▶ ▲目▶ ▲目▶ ▲□▶