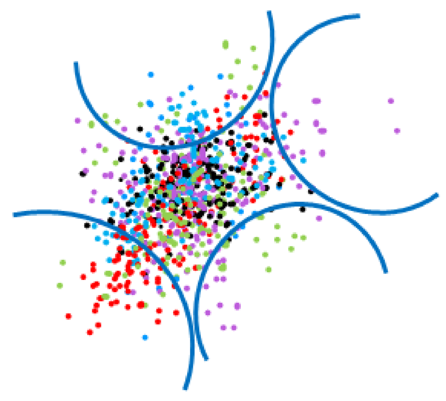
GuCNet: A Guided Clustering-based Network for Improved Classification

Ushasi Chaudhuri, Syomantak Chaudhuri, Subhasis Chaudhuri

Indian Institute of Technology Bombay, India.

Introduction to Problem

- **Task:** 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.
- **Problem:** Semantic classification of challenging and highly-cluttered data is difficult.

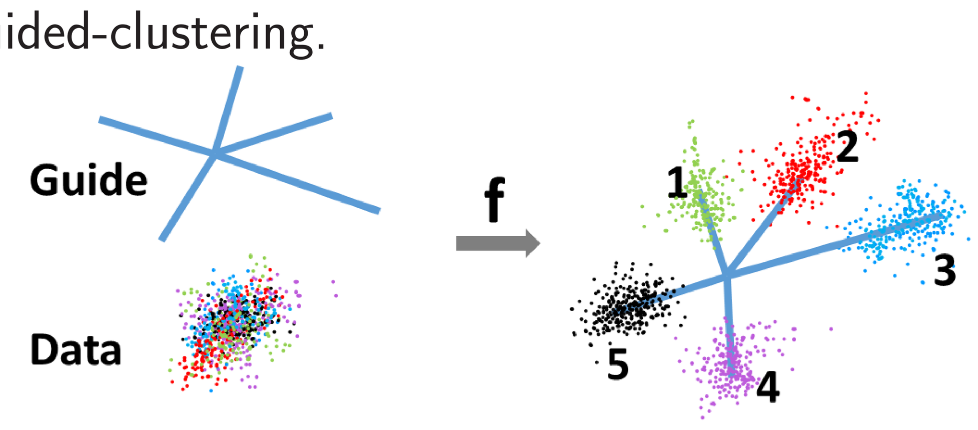


Guided Clustering

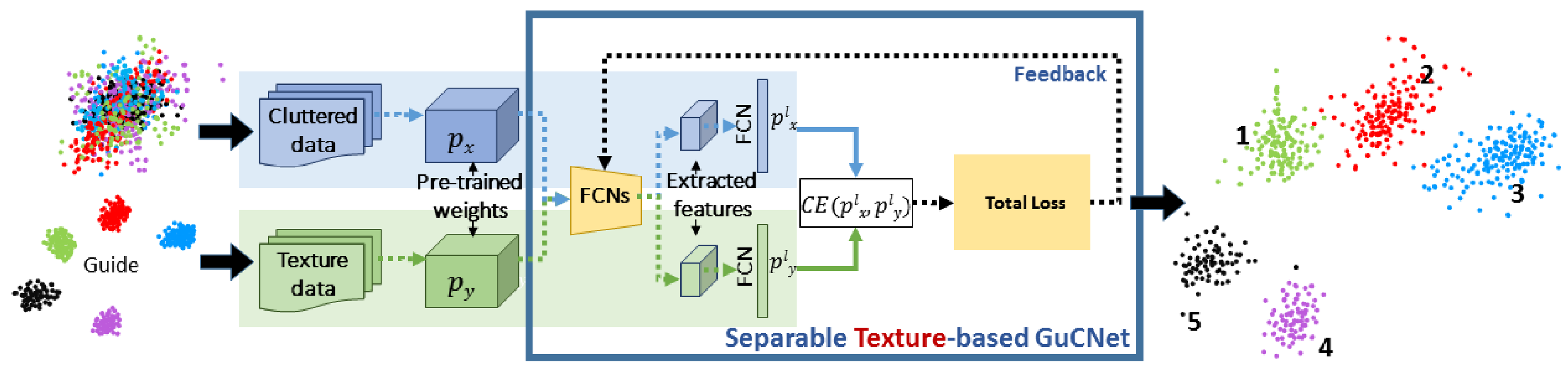
Many well-separable datasets are available.

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

- **Guide data ($\mathcal{X}$):** A well separable data.
- **Cluttered data ($\mathcal{Y}$):** 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.



Texture-based Guiding

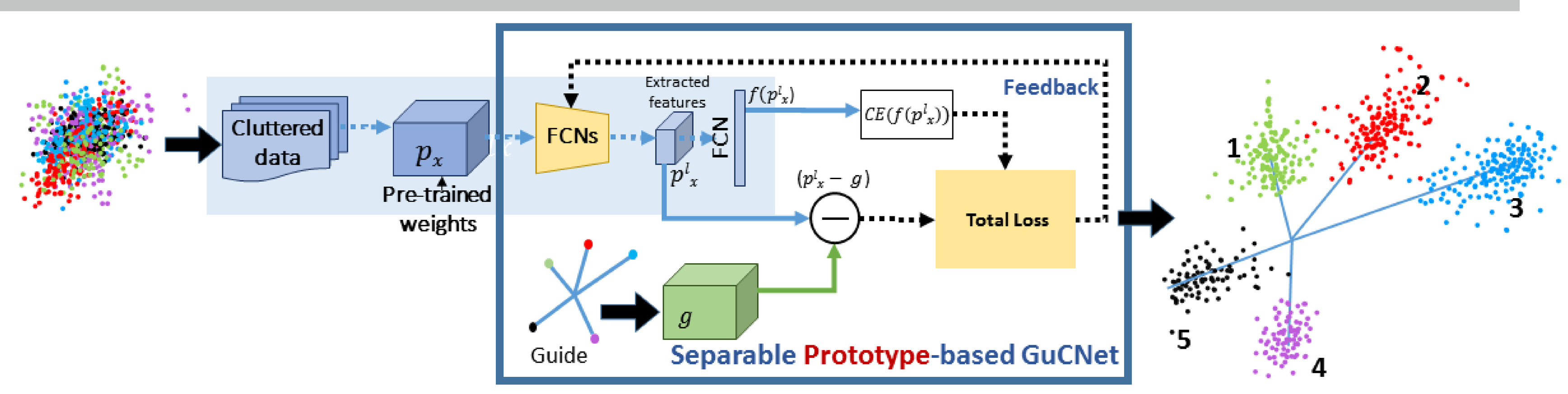


A well-separable data acts as a texture data.

We refer to the output of the convoluted features from $\mathcal{X}$ and $\mathcal{Y}$ as p^l_x and p^l_x, respectively.

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

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 ($K \geq C$) dimensional vectors (called *prototypes* g).
- **Matching loss:** ($\mathcal{L}_{\text{ml}} = \mid p^l_x - g \mid$) to bring the dataset closer to the prototype vectors.
- Minimize $\mathcal{L}_{\text{ml}} = \mid p^l_x - g \mid + \text{CE}(p^l_x, p^l_y)$.

Results - RSSCN aerial scene dataset

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

Model	Accuracy(%)
VRGSIR	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

Classification performance on LSUN dataset with the same guide data.

Model	Accuracy(%)
Vanilla GAN	70.50%
Labeled-samples	77.00%
Hybrid GAN	83.20%
Normal BN-Inception	90.40%
Deeper BN-Inception	90.90%
SJTU-ReadSense	90.40%
SIAT MMLAB	91.60%
Baseline	83.75%
GuCNet (Prototype)	**95.03%**
GuCNet (Texture)	**94.86%**

Results - TU-Berlin sketch dataset

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%
Alexnet-FC-GRU	79.95%
Zhang *et. al.*	82.95%
Baseline	69.90%
GuCNet (Prototype)	**86.63%**
GuCNet (Texture)	**89.26%**

Some Interesting Ablation Study:

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

Dataset (TU-Berlin)	**Accuracy(%)**
Same class binning	89.26%
Dissimilar class binning	90.05%

Effect of separability of prototypes in terms of Hamming distance (H).

Dataset	**Separation of prototypes**			
	w2vec	$H=2$	$\frac{H_{\text{max}}}{2}$	H_{max}
RSCCN	96.20%	96.02%	96.27%	97.36%
LSUN	92.71%	94.60%	94.92%	95.03%

Conclusions

- Propose a simple guided clustering framework 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.