Deep Self-Learning From Noisy Labels

Presented by: Ashkan Farhangi

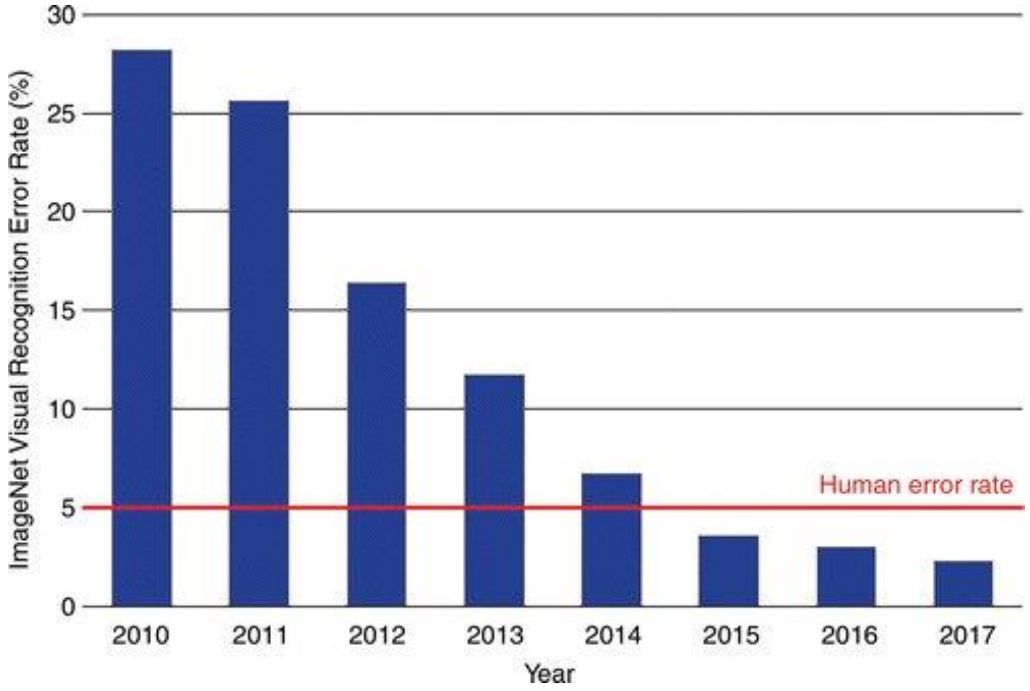
Jiangfan Han
Ping Luo
Xiaogang Wang

Outline

- Introduction
- Related works
- Proposed method
- Experiments & results

History: ImageNet

- Large-scale well-labeled training data.
- 14 million images hand-annotated by the project.
- Neural Networks achieve better performance than humans.
- Quite expensive data labeling process.



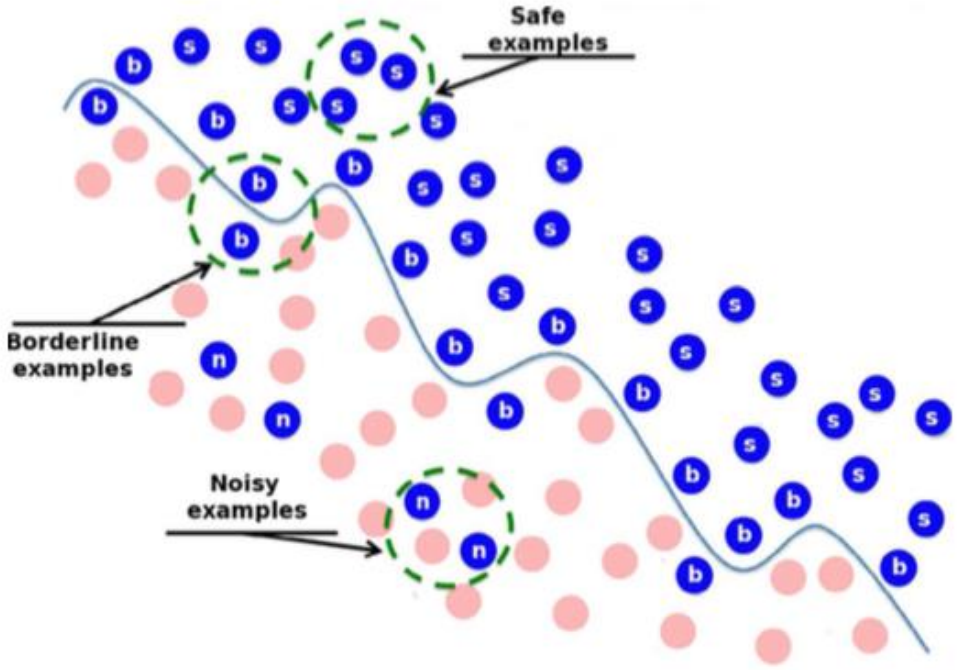


http://www.image-net.org/

Solution

- Supervised and unsupervised deep learning algorithms

Dog





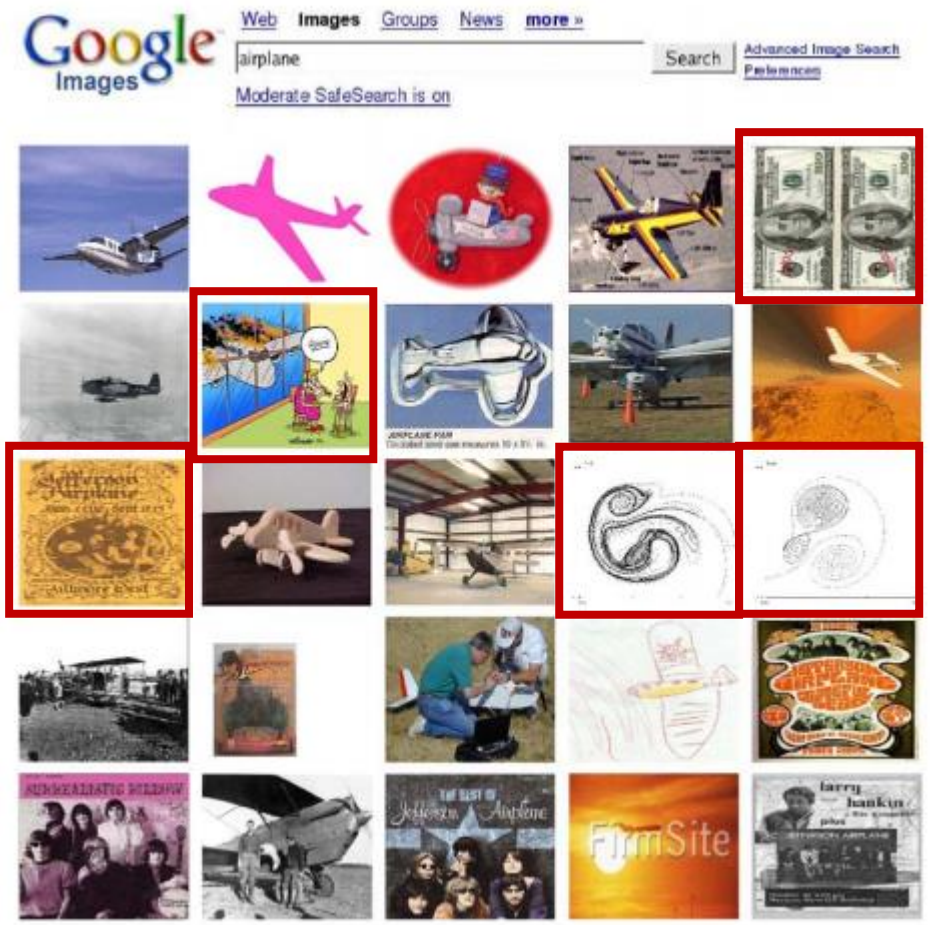
Muffin

https://sci2s.ugr.es/noisydata

https://www.freecodecamp.org/news/chihuahua-or-muffin-my-search-for-the-best-computer-vision-api-cbda4d6b425d/

Solution

- Automatically collect noisy label data from Internet.



Noisy label

clean cat



noisy cat



clean jasmine



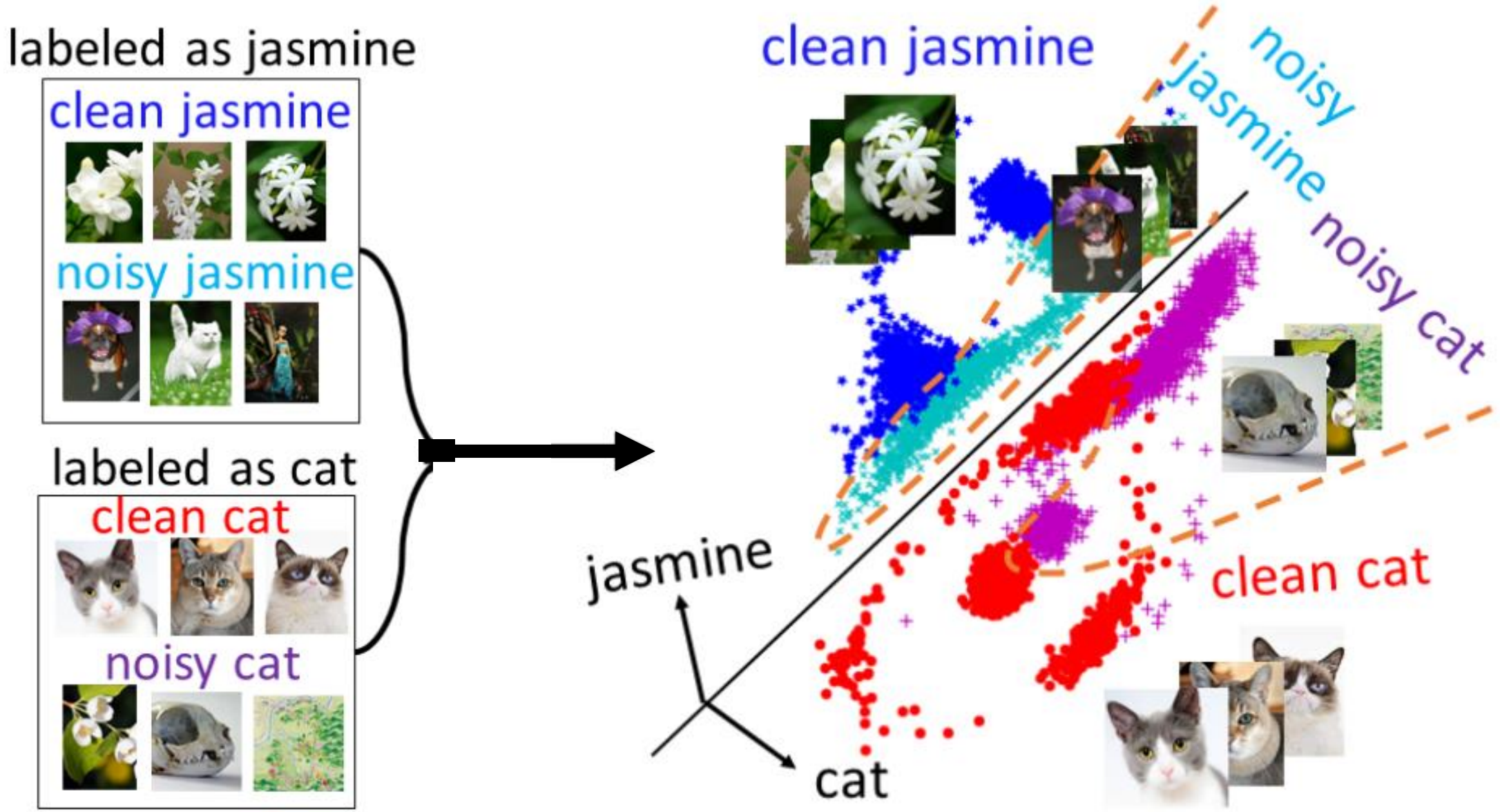
noisy jasmine



Noisy label challenges

- Researchers need to develop algorithms for learning in presence of label noise.
- Human supervision for label correction is costly but effective.
- Approaches not relying on human supervision are scalable but less effective.
- There is a need to automate the label correction process.

Distribution of features



Wang, Y., Liu, W., Ma, X., Bailey, J., Zha, H., Song, L. and Xia, S.T., 2018. Iterative learning with open-set noisy labels. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* (pp. 8688-8696).

Existing solutions

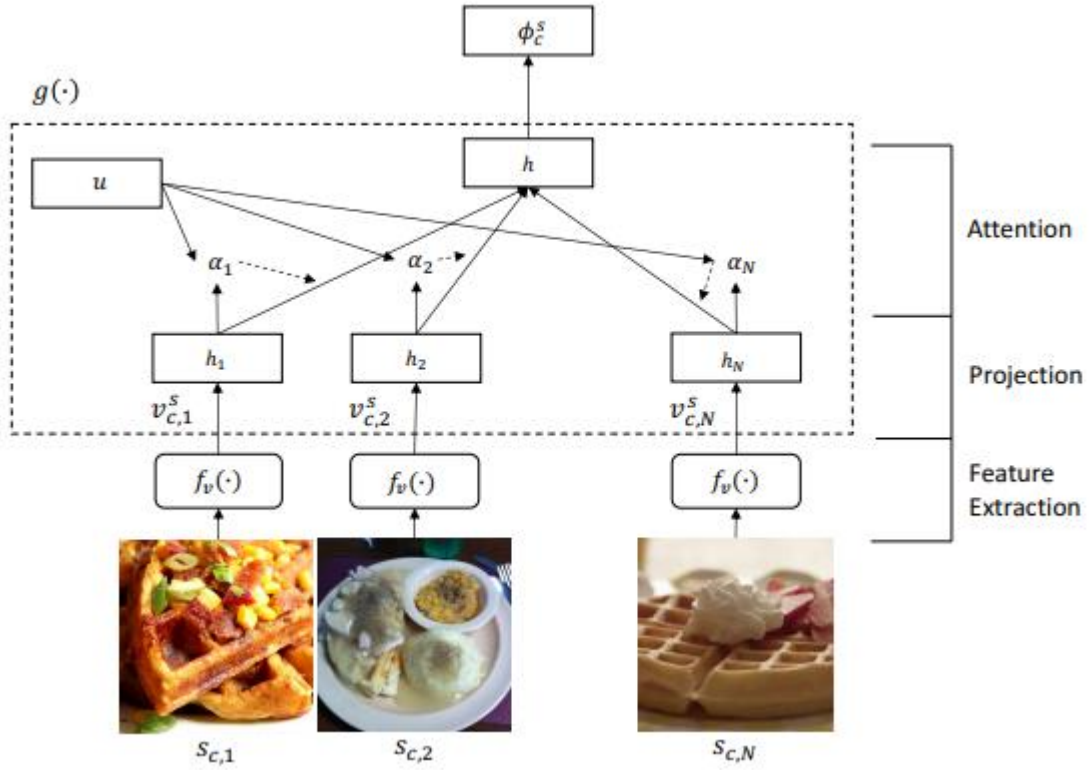
- **Supervised:** Supervised methods learn a binary classification from verification labels for each class.
  - CleanNet, 2-layer MLP, kNN, SVM
- **Unsupervised:** Unsupervised methods rely on the outlier removal.
  - DRAE
- **Weakly supervised:** verification labels are not required.
  - Classification filtering, Loss Factorization

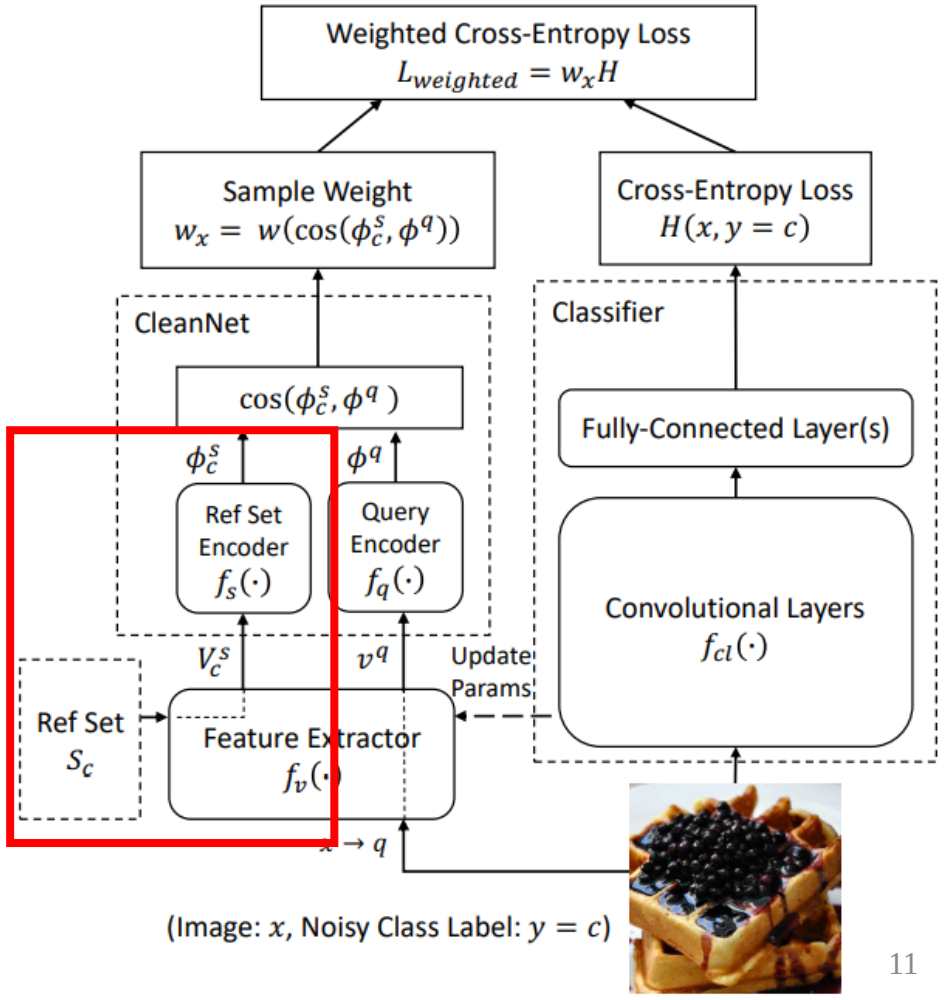
Related Work – CleanNet

- Utilizes two networks in one framework for end-to-end learning of image classifiers with label noise
- Uses CNN- image classifier to provides more discriminative features.
- It also uses attention mechanism to assign weights to data samples

Related Work – CleanNet

- CleanNet uses the concept of class prototype to represent each class category





Related Work – CleanNet

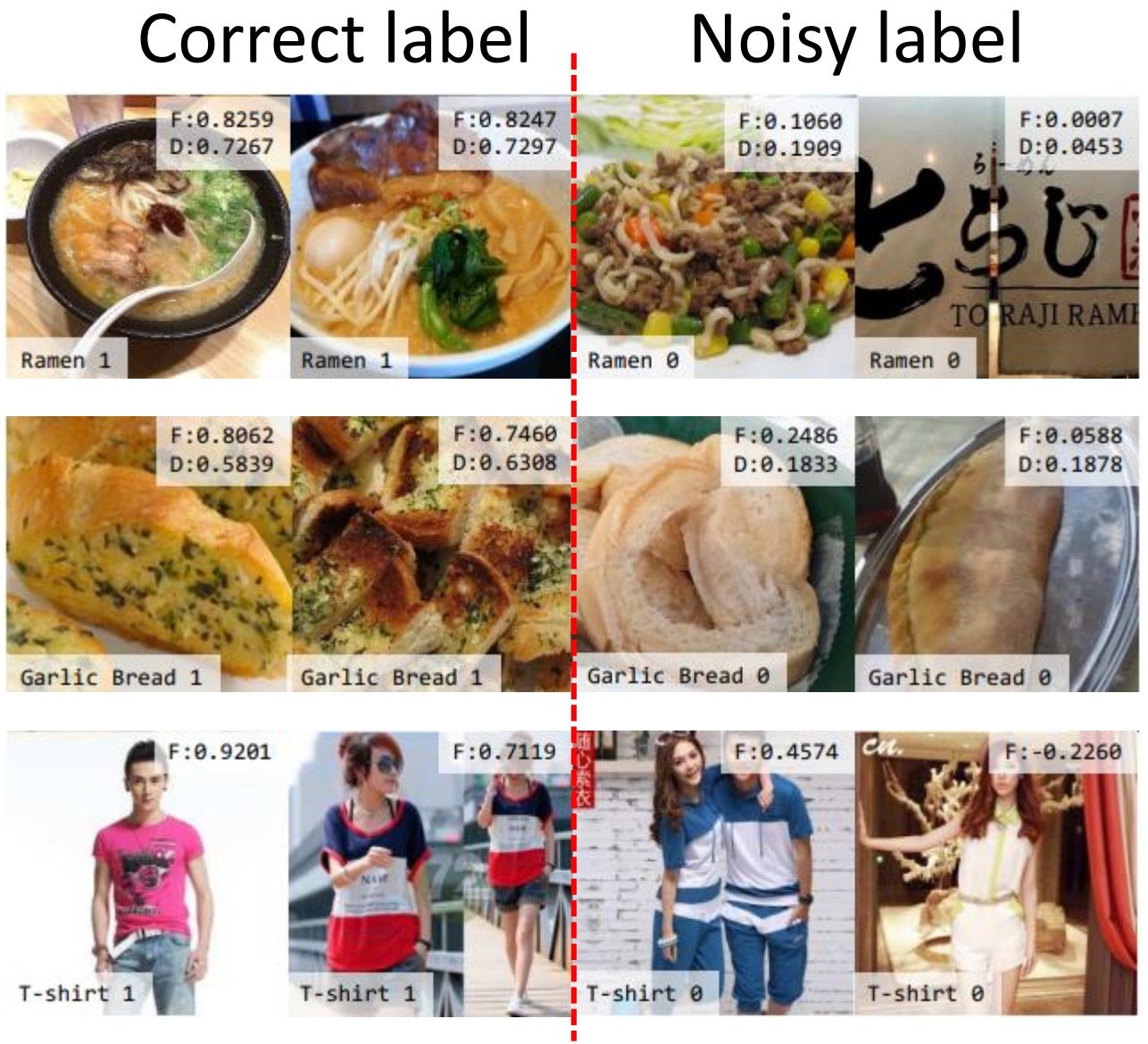
- Class prototypes can effectively represent classes.
- Compare the sample image with the prototype will determine if the label is correct.



$$u_i = \tanh(Wh_i + b)$$

$$\alpha_i = \frac{\exp(u_i^T u)}{\sum_i \exp(u_i^T u)}$$

Related Work – CleanNet



Lee, K.H., He, X., Zhang, L. and Yang, L., 2018. Cleannet: Transfer learning for scalable image classifier training with label noise. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* (pp. 5447-5456).

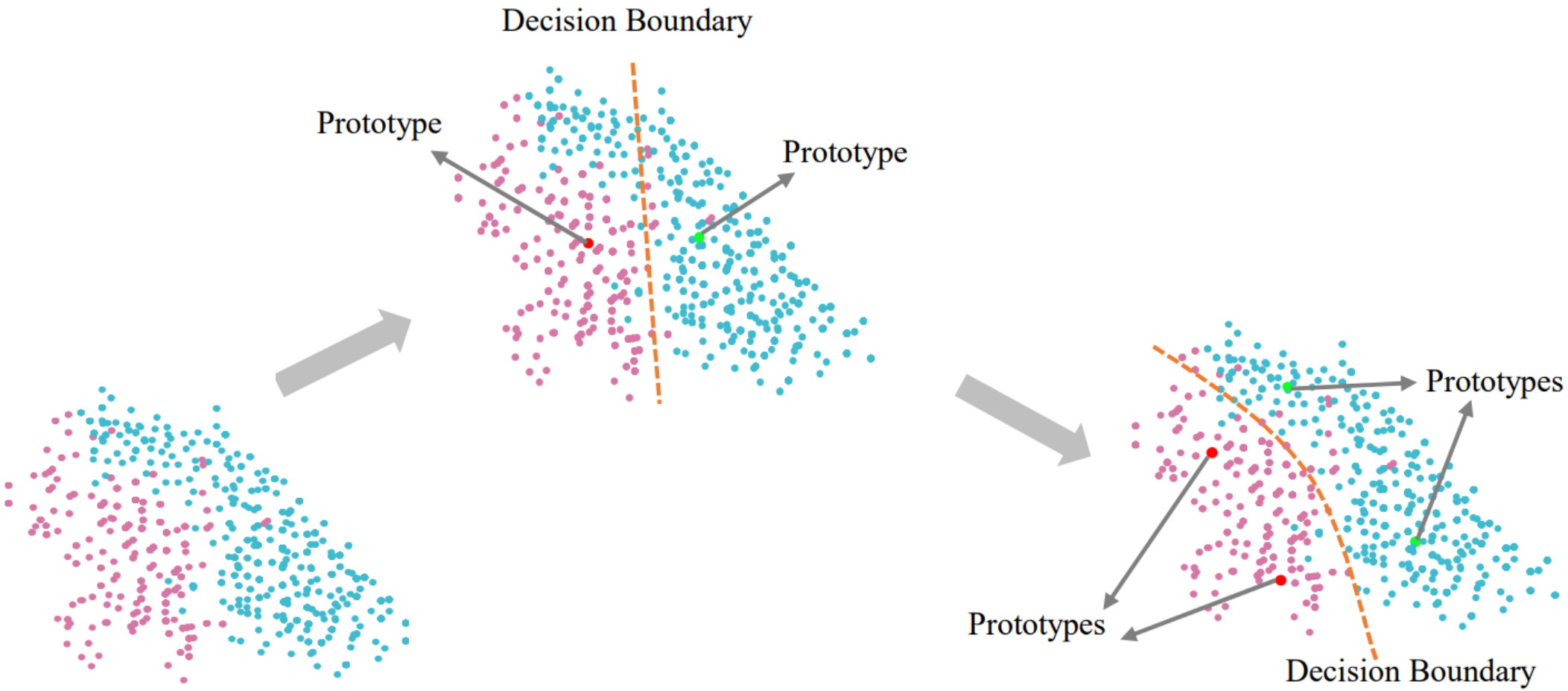
Related work

- Majority of these approaches require extra clean samples.
- Using a complicated training procedure.
- Relying on the noisy dataset that is created intentionally by human with label flipping.
  - CIFAR10

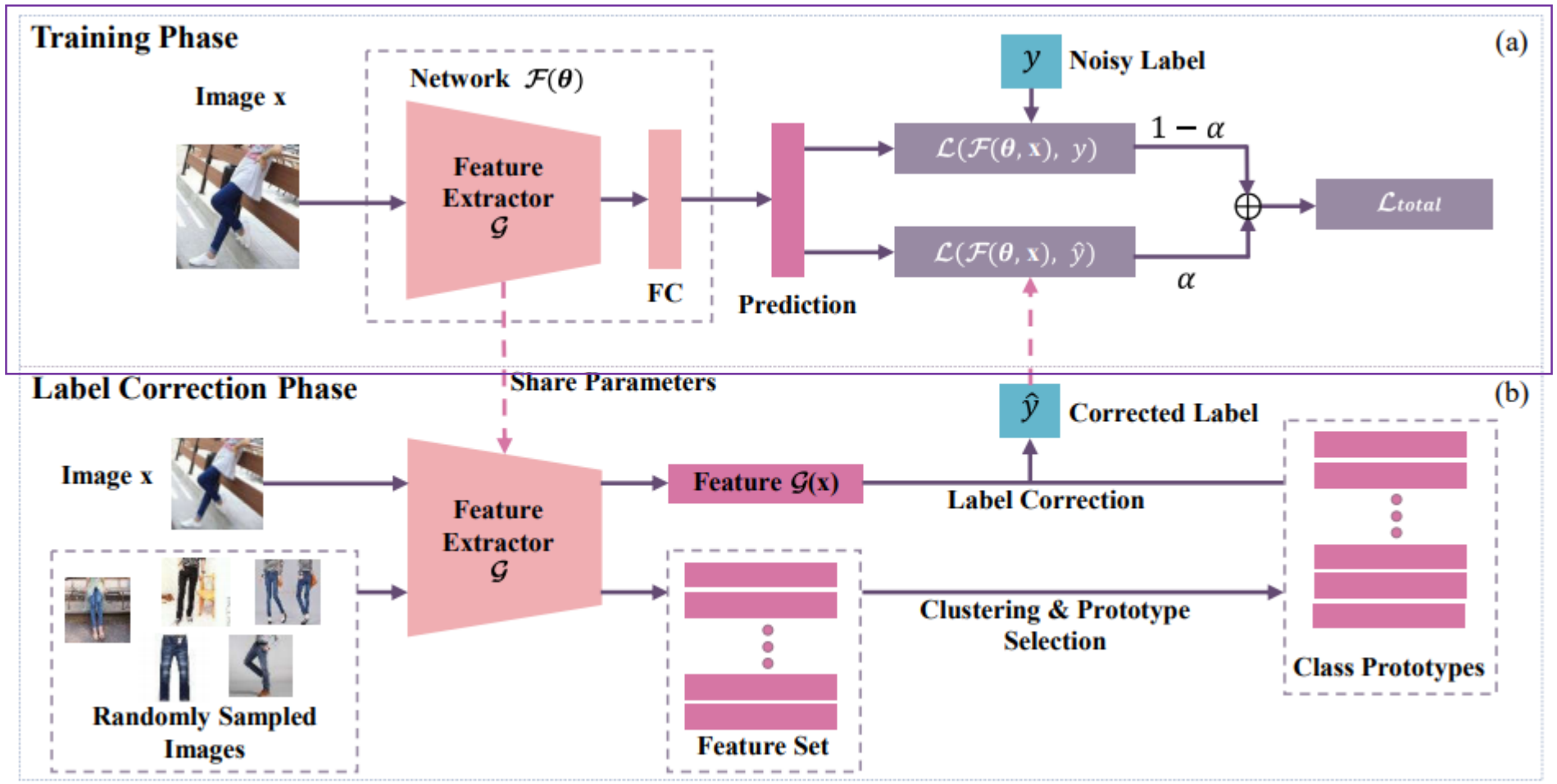
Deep Self-Learning for noisy labels

- Select images as prototypes by their density and similarity according to the data distribution.
- Not rely on clean supervision images.
- Optimal number of prototypes for each class.

Deep Self-Learning for noisy labels

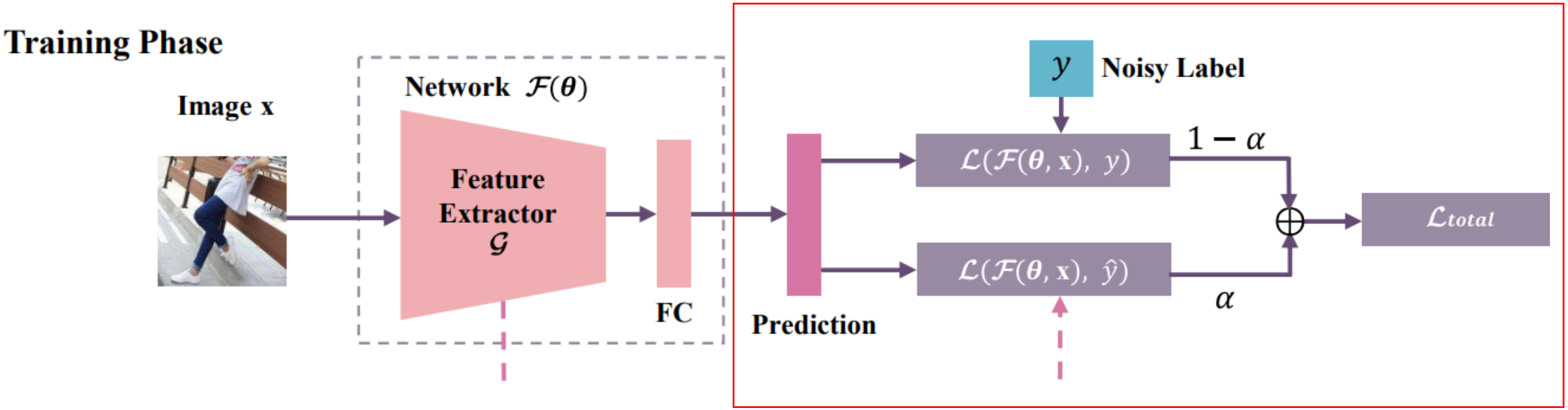


Proposed network



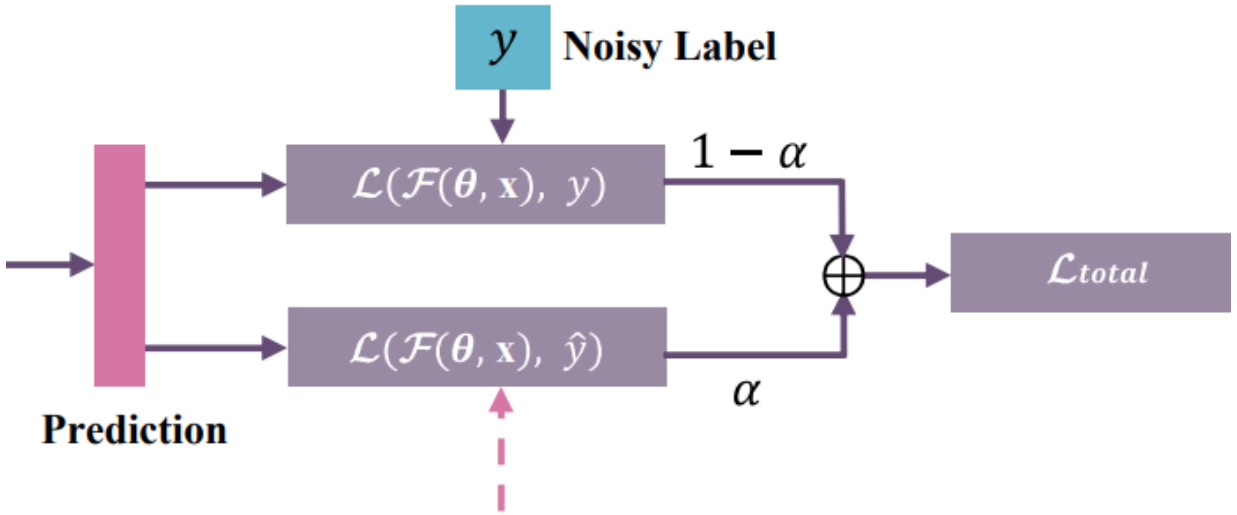
Training Phase

$$\mathcal{L}(\mathcal{F}(\theta, \mathbf{x}), y) = -\frac{1}{n}\sum_{i=1}^{n}\log\left(\mathcal{F}(\theta, \mathbf{x}_i)_{y_i}\right)$$



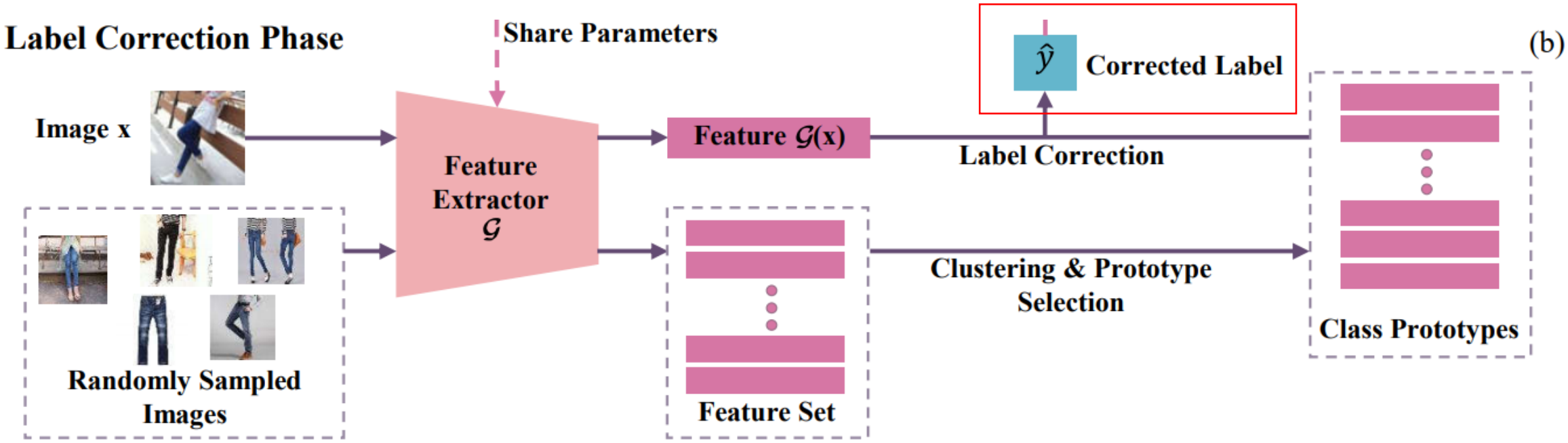
Training Phase Losses

$$\mathcal{L}(\mathcal{F}(\theta, \mathbf{x}), y) = -\frac{1}{n}\sum_{i=1}^{n} \log\left(\mathcal{F}(\theta, \mathbf{x}_i)_{y_i}\right)$$



$$\mathcal{L}_{\text{total}} = (1-\alpha)\mathcal{L}(\mathcal{F}(\theta, \mathbf{x}), y) + \alpha\mathcal{L}(\mathcal{F}(\theta, \mathbf{x}), \hat{y})$$

Label Correction Phase

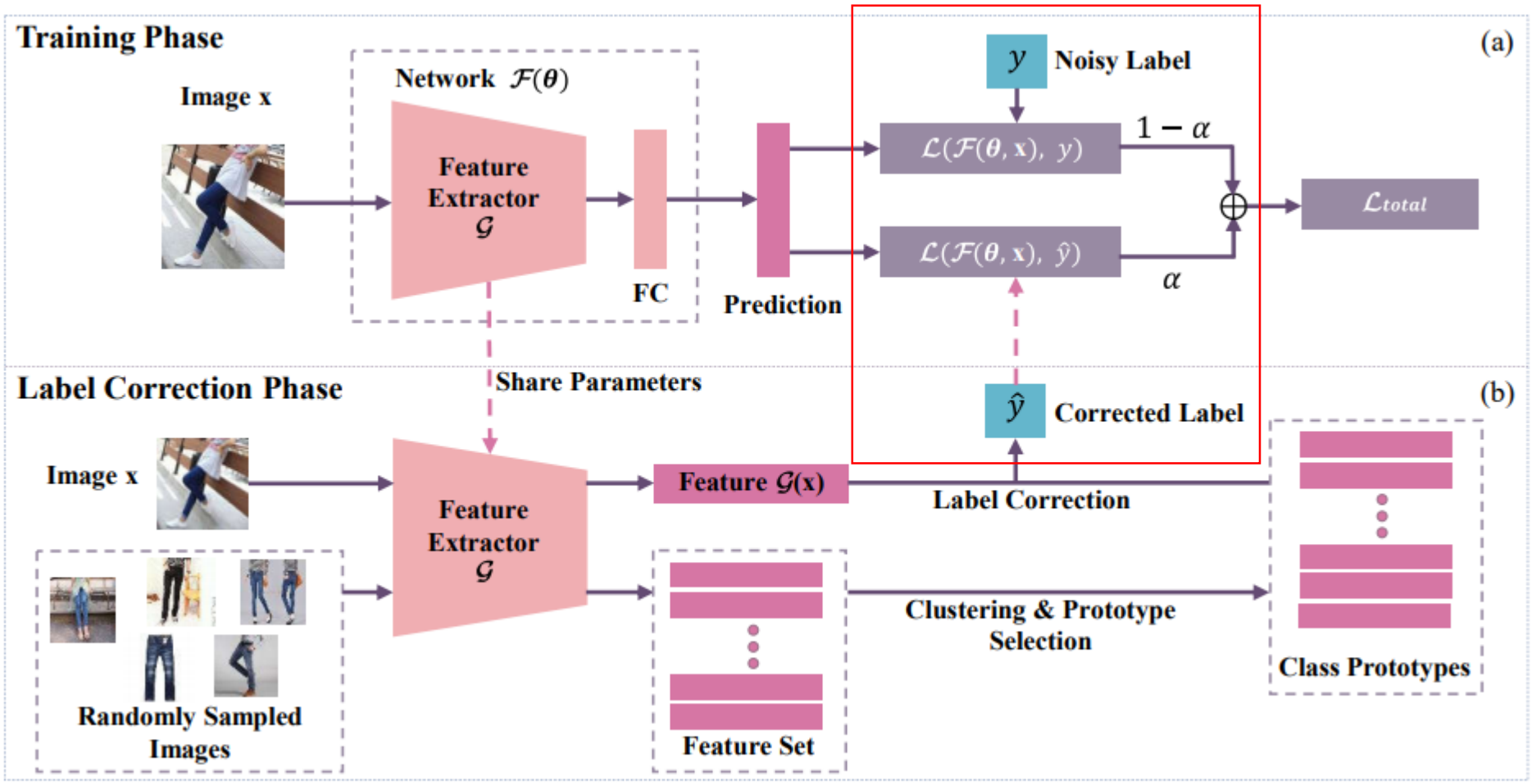


$$\sigma_c = \frac{1}{p}\sum_{l=1}^{p} \cos(\mathcal{G}(\mathbf{x}), \mathcal{G}(\mathbf{x}_{cl})), c = 1...K$$

$$\hat{y} = \text{argmax}_c\, \sigma_c, \quad c = 1...K$$

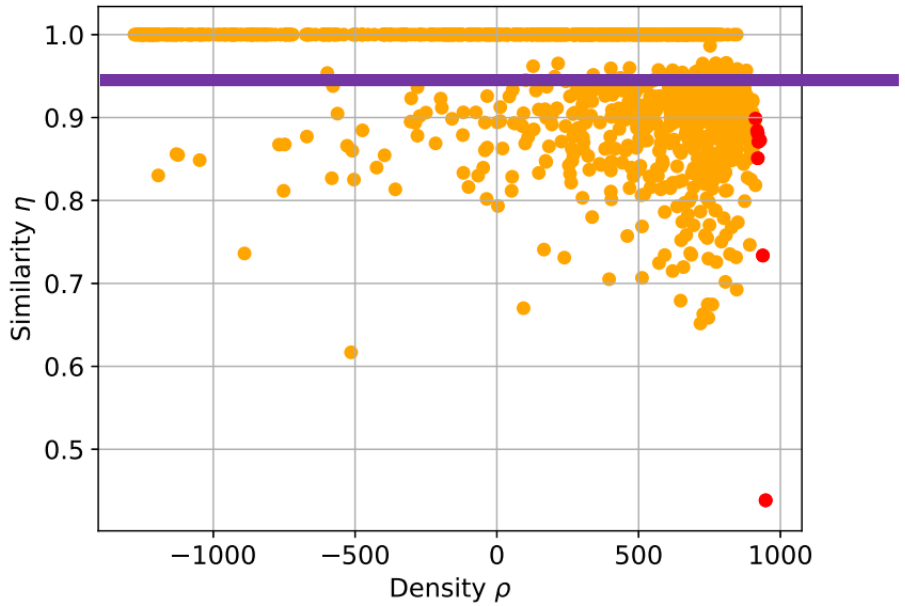
Proposed network

$$\mathcal{L}_{\text{total}} = (1-\alpha)\mathcal{L}(\mathcal{F}(\theta, \mathbf{x}), y) + \alpha\mathcal{L}(\mathcal{F}(\theta, \mathbf{x}), \hat{y})$$



Distribution

- Over 80% of the samples have η > 0.9
- Half of the samples have η > 0.95.
- high-density value ρ and low similarity value η can be chosen as a prototype.



Food101-N Dataset

- The Food-101N dataset was introduced in the CleanNet paper.
- 310,009 images of food recipes classified in 101 classes (categories).





Source: https://kuanghuei.github.io/Food-101N/

Clothing1M Dataset

- It consists of 1M images with noisy class labels from 14 fashion classes



https://github.com/Cysu/noisy_label

Experiment results

#	Method	Data	Accuracy
1	Cross Entropy	1M noisy	69.54
2	Forward [25]	1M noisy	69.84
3	Joint Optim. [35]	1M noisy	72.23
4	MLNT-Teacher [16]	1M noisy	73.47
5	Ours	1M noisy	**74.45**
6	Forward [25]	1M noisy + 25k verify	73.11
7	CleanNet w_{hard} [15]	1M noisy + 25k verify	74.15
8	CleanNet w_{soft} [15]	1M noisy + 25k verify	74.69
9	Ours	1M noisy + 25k verify	**76.44**
10	Cross Entropy	1M noisy+ 50k clean	80.27
11	Forward [25]	1M noisy + 50k clean	80.38
12	CleanNet w_{soft} [15]	1M noisy + 50k clean	79.90
13	Ours	1M noisy + 50k clean	**81.16**

Clothing1M

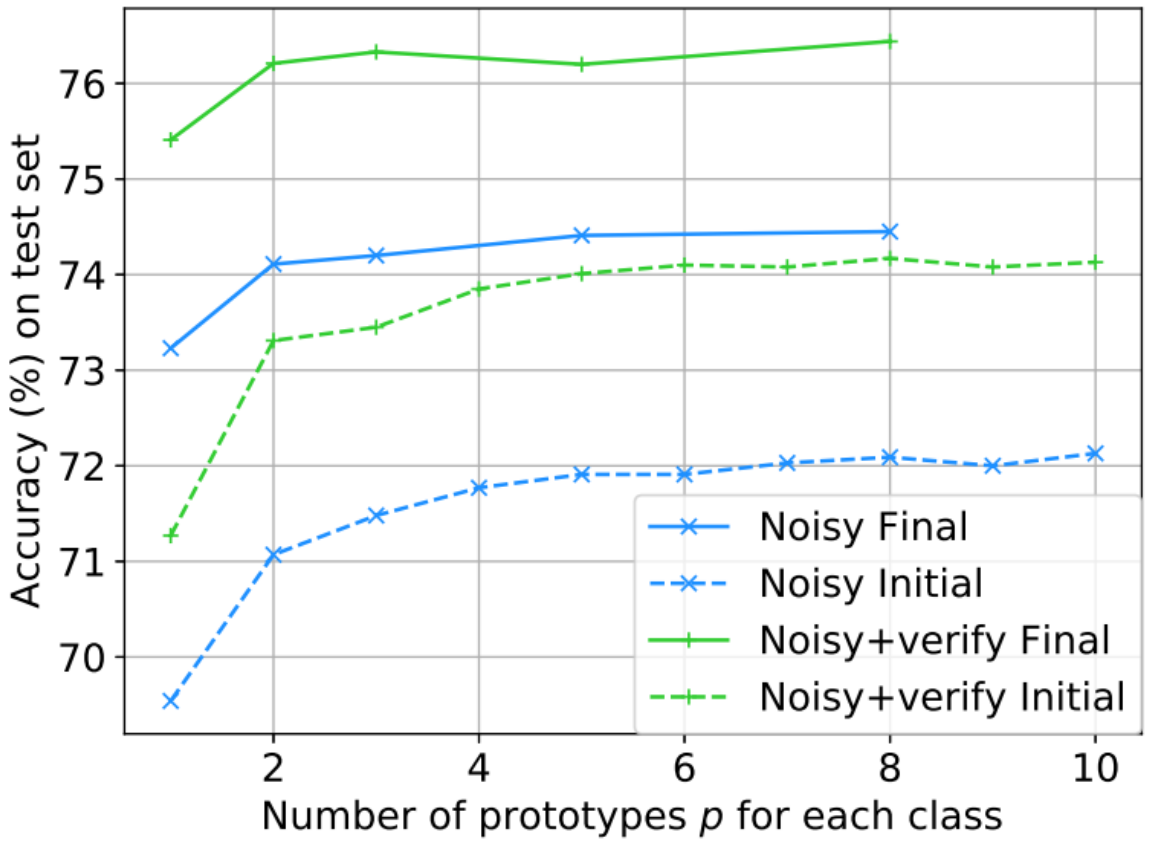
#	Method	Accuracy
1	Cross Entropy	84.51
2	CleanNet w_{hard} [15]	83.47
3	CleanNet w_{soft} [15]	83.95
4	Ours	**85.11**

Food-101N

Prototypes per class

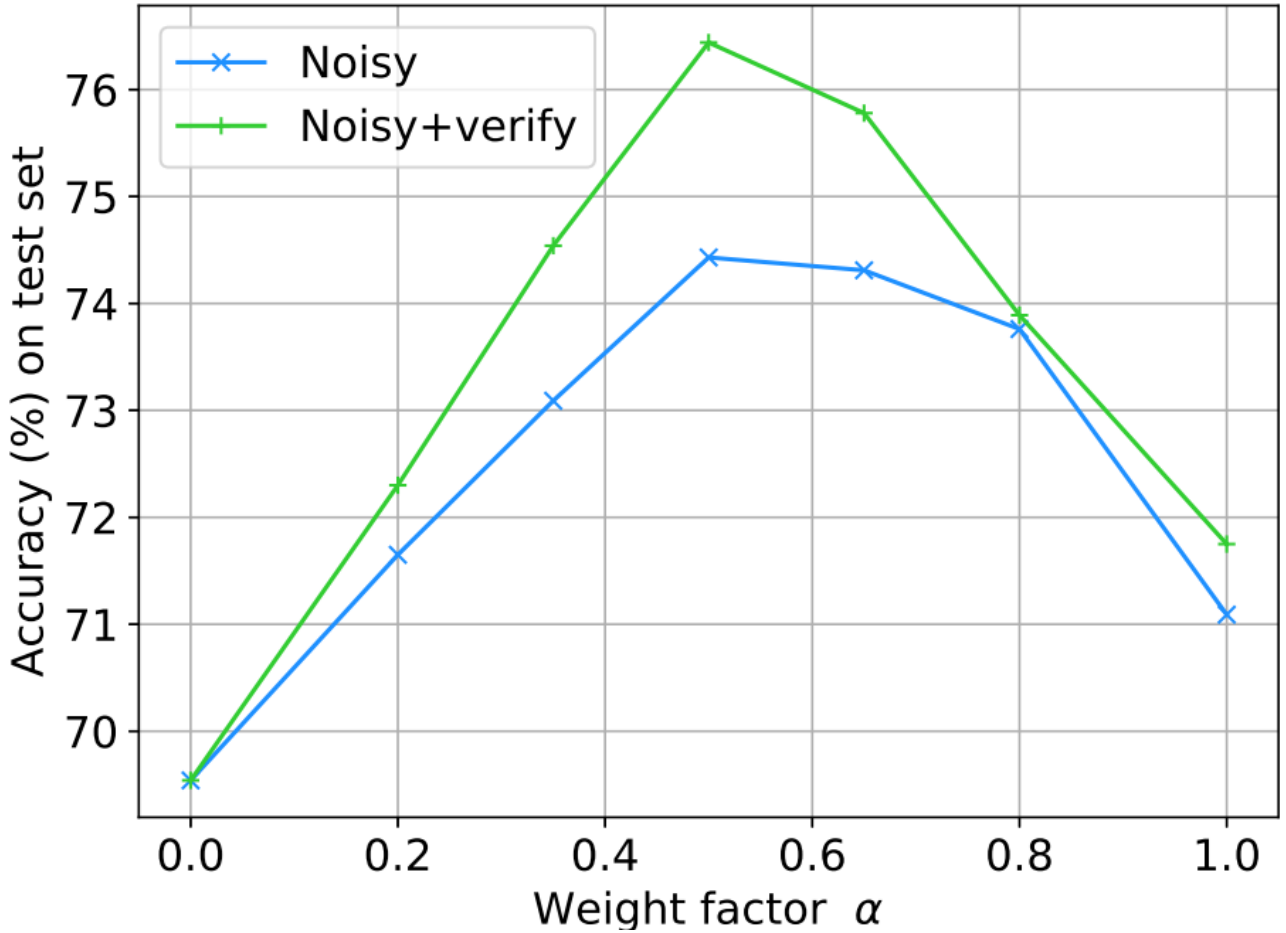
- The number of class prototypes is the key to the representation ability to a class

$$\mathcal{L}_{\text{total}} = (1-\alpha)\mathcal{L}(\mathcal{F}(\theta, \mathbf{x}), y) + \alpha\mathcal{L}(\mathcal{F}(\theta, \mathbf{x}), \hat{y})$$



Weight factor

- Decides the network will concentration on the noisy labels and corrected labels.



Deep self-learning method – Food101-N



Edamame Dumpling



Apple pie Cup cake



Churros Ice cream



Sashimi Sushi

Deep self-learning method – Clothing1M



Hoodie	Jacket



Shirt	T-Shirt

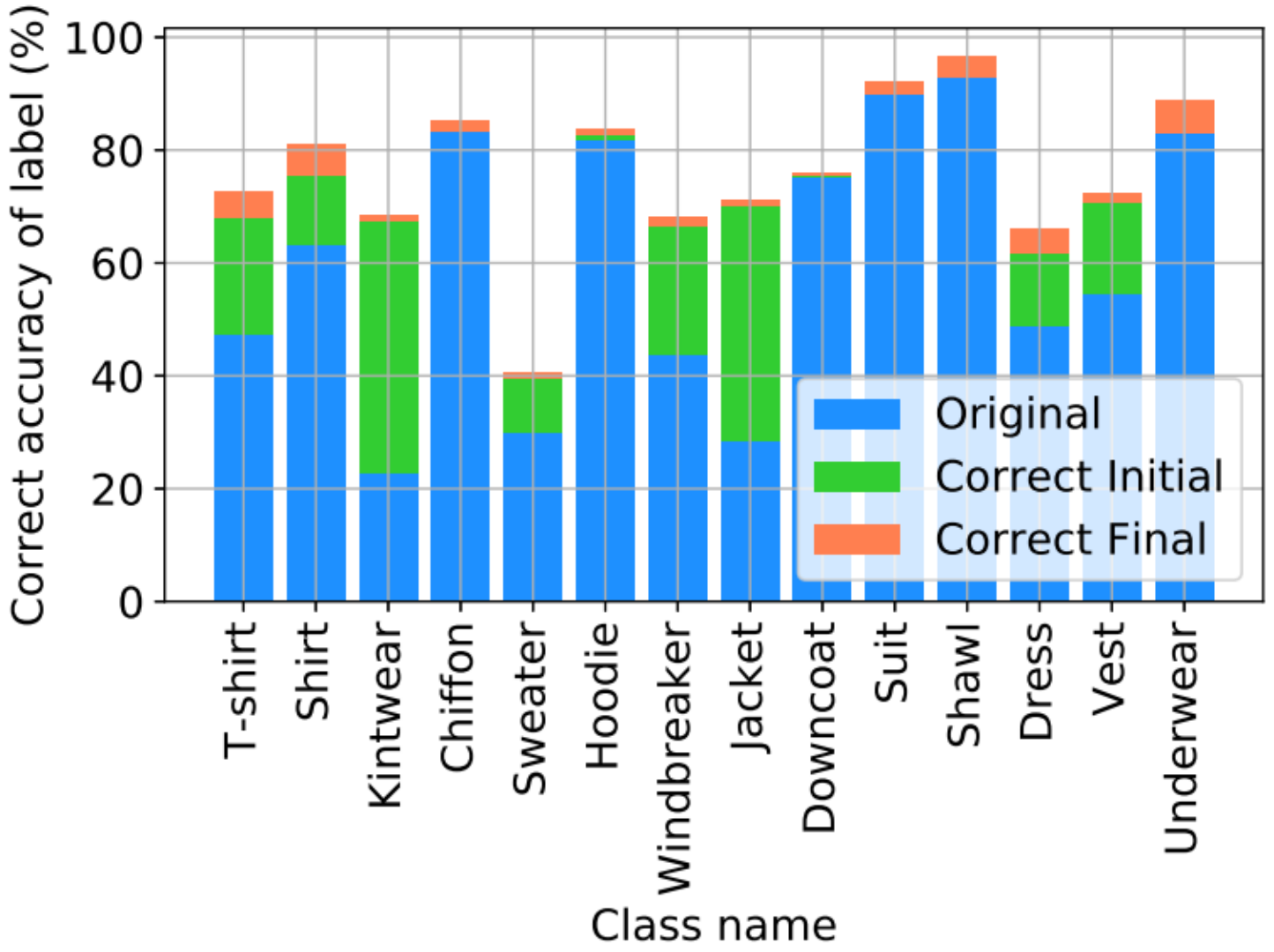


Jacket	Hoodie



Jacket	Suit

Insights from hard classes



Cluster methods

- The classification accuracy with various clustering methods
- Further research can be done in this area to find prototypes.

Method	Data	Accuracy
K-means++ [1]	1M noisy	74.08
Density peak Euc. [31]	1M noisy	74.11
Ours	1M noisy	**74.45**
K-means++ [1]	1M noisy + 25k verify	76.22
Density peak Euc. [31]	1M noisy + 25k verify	76.05
Ours	1M noisy + 25k verify	**76.44**

Sources:

- Han, J., Luo, P. and Wang, X., 2019. Deep self-learning from noisy labels. In *Proceedings of the IEEE International Conference on Computer Vision* (pp. 5138-5147).
- Lee, K.H., He, X., Zhang, L. and Yang, L., 2018. Cleannet: Transfer learning for scalable image classifier training with label noise. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* (pp. 5447-5456).
- Wang, Y., Liu, W., Ma, X., Bailey, J., Zha, H., Song, L. and Xia, S.T., 2018. Iterative learning with open-set noisy labels. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* (pp. 8688-8696).

Thank you.

Questions?