Open DMQA Seminar

Noisy Label Learning

김상훈

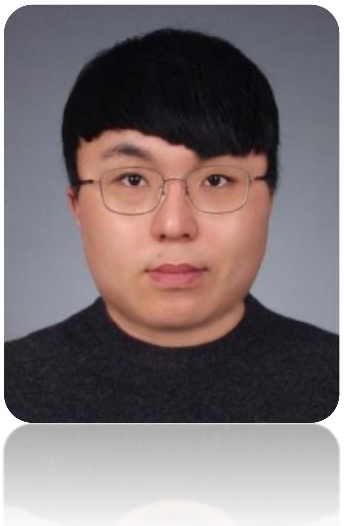
Data Mining & Quality Analytics Lab.

2023.06.22

고려대학교 KOREA UNIVERSITY

Data Mining Quality Analytics

발표자 소개



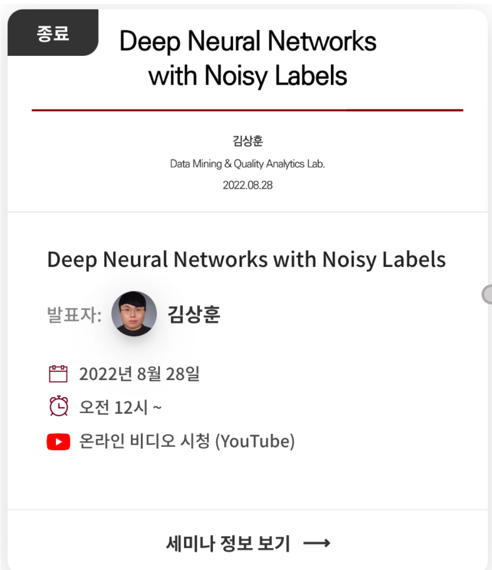
- **김상훈 (Sanghoon Kim)**
  - ✓ 고려대학교 산업경영공학과
  - ✓ Data Mining & Quality Analytics Lab. (김성범 교수님)
  - ✓ Ph.D. Student (2019.09 ~ Present)
- **Research Interest**
  - ✓ Robust AI for real-world dataset
    - ✓ Open-set recognition
    - ✓ Noisy Label Learning
- **Contact**
  - ✓ E-mail : dawonksh@korea.ac.kr

목차

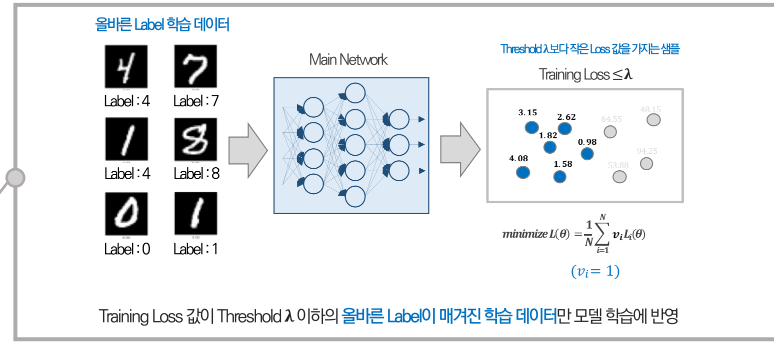
Noisy Label Learning

지난 세미나 :

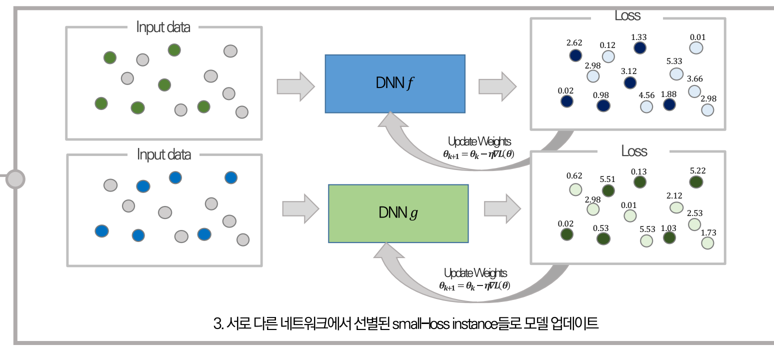
http://dmqa.korea.ac.kr/activity/seminar/377



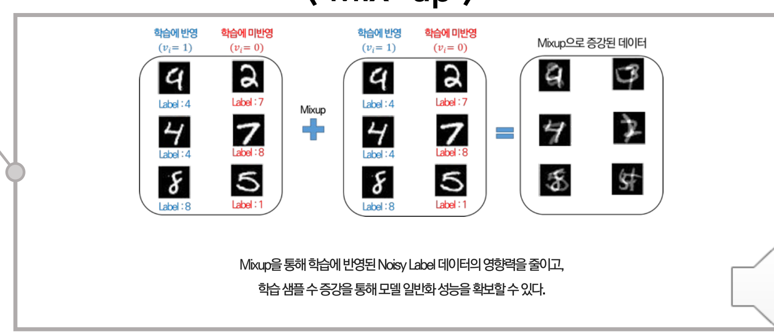
〈 Loss value 〉



〈 Co-learning 〉



〈 Mix-up 〉



목차

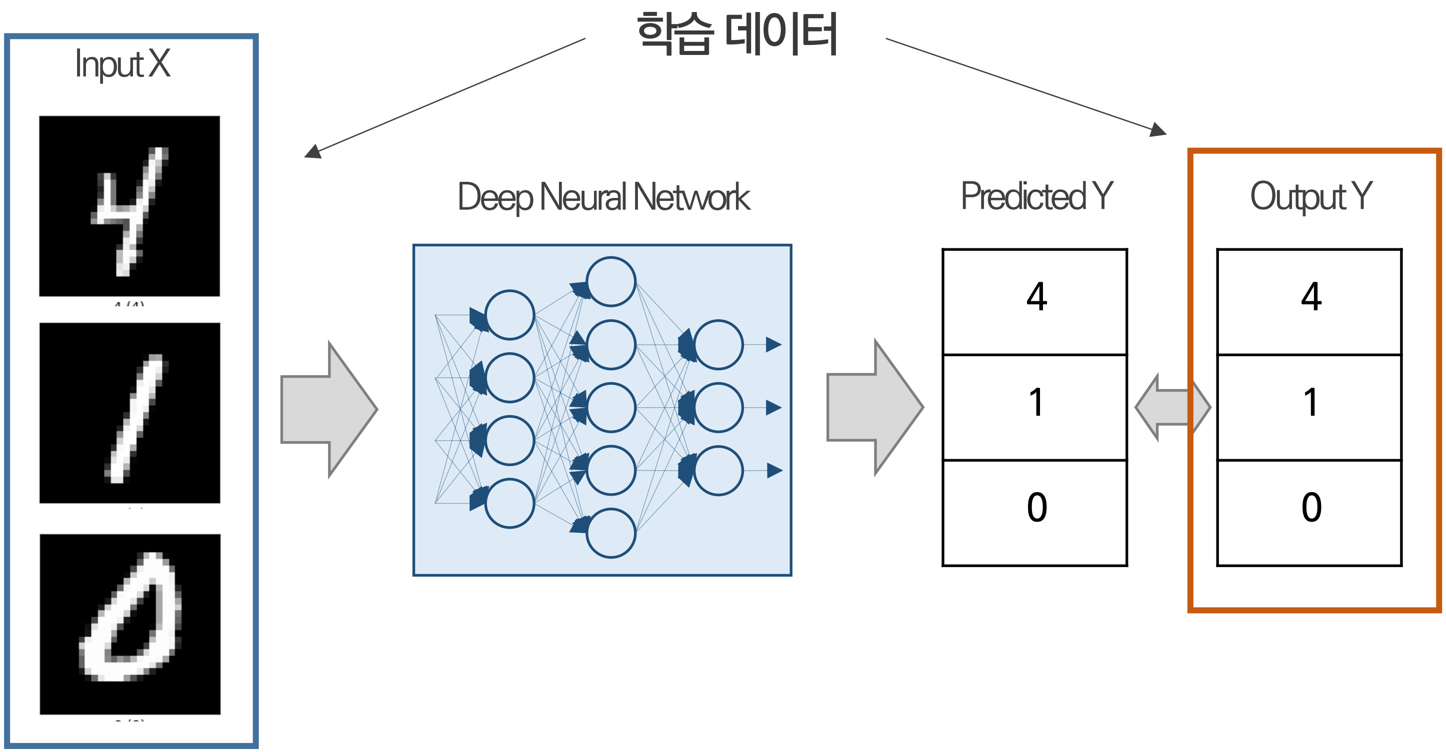
Noisy Label Learning

- **Noise Transition Matrix**
  - Hendrycks, D., Mazeika, M., Wilson, D., & Gimpel, K. (2018). Using trusted data to train deep networks on labels corrupted by severe noise. Advances in neural information processing systems, 31.
  - Goldberger, J., & Ben-Reuven, E. (2017, April). Training deep neural-networks using a noise adaptation layer. In In ternational conference on learning representations
- **Robust Loss Regularization**
  - ELR : Liu, S., Niles-Weed, J., Razavian, N., & Fernandez-Granda, C. (2020). Early-learning regularization preve nts memorization of noisy labels. Advances in neural information processing systems, 33, 20331-20342.

Noise Transition Matrix

❖ 일반적으로 Supervised Learning에서는 학습데이터 모두 정확한 Label이 붙어있음을 가정

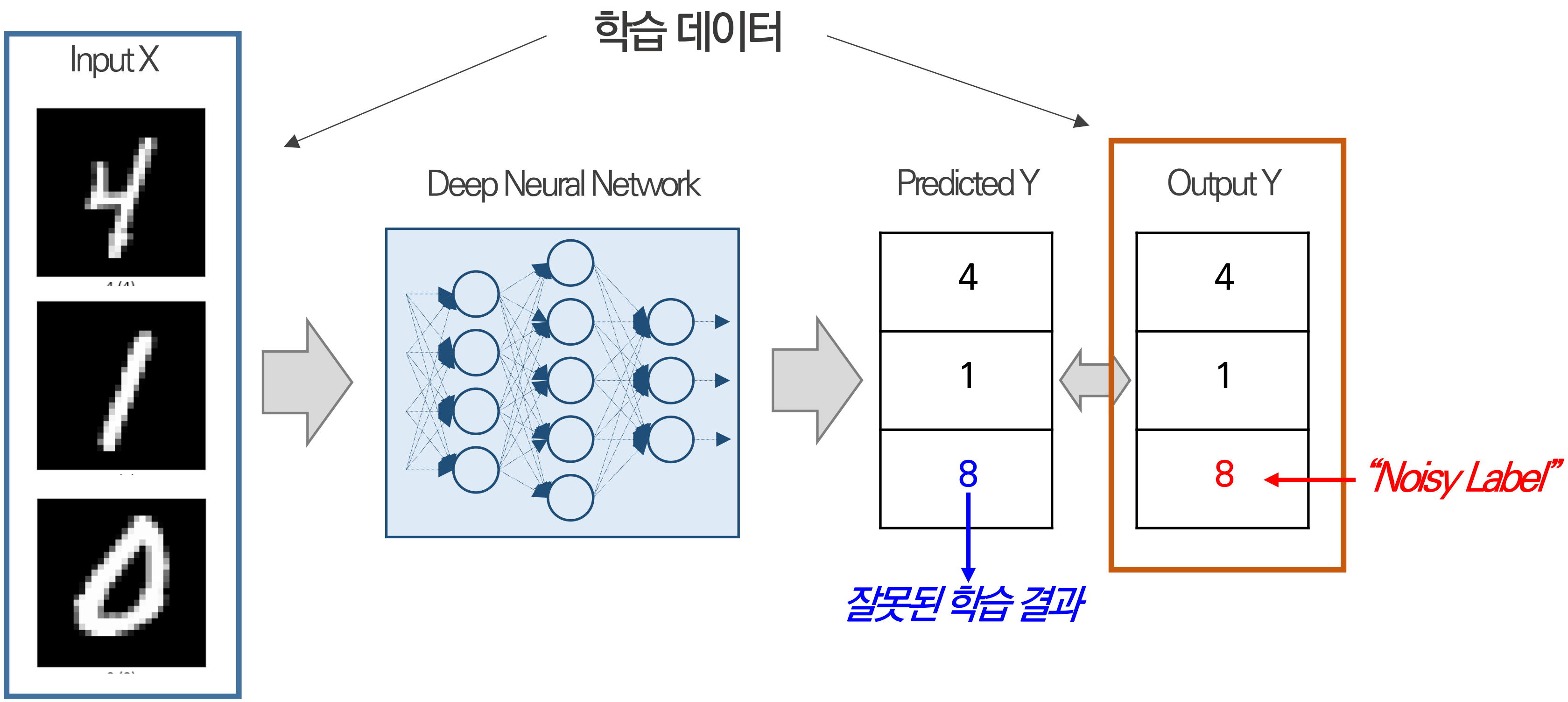
- Labeling 작업이 모두 완벽하게 이루어진 정제된 실험 데이터를 이용하여 입력 이미지와 레이블의 관계를 모델이 학습



Noise Transition Matrix

❖ 일반적으로 Supervised Learning에서는 학습데이터 모두 정확한 Label이 붙어있음을 가정

- Labeling 작업이 모두 완벽하게 이루어진 정제된 실험 데이터를 이용하여 입력 이미지와 레이블의 관계를 모델이 학습

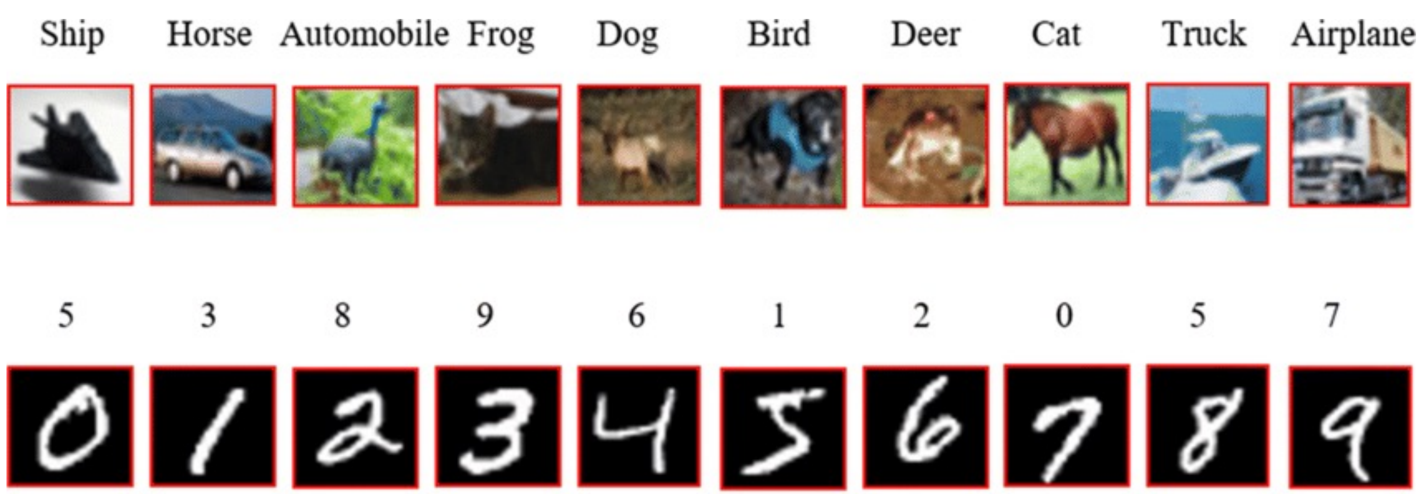


Noisy Labels는 Deep Learning 모델의 성능을 저하시키는 주요 원인

Noise Transition Matrix

❖ Noisy Label의 종류

- **Random Noise:** Instance Feature이나 Class에 영향을 받지 않음
  - 아무런 규칙없이 일정 확률로 잘못된 Label을 가지는 데이터



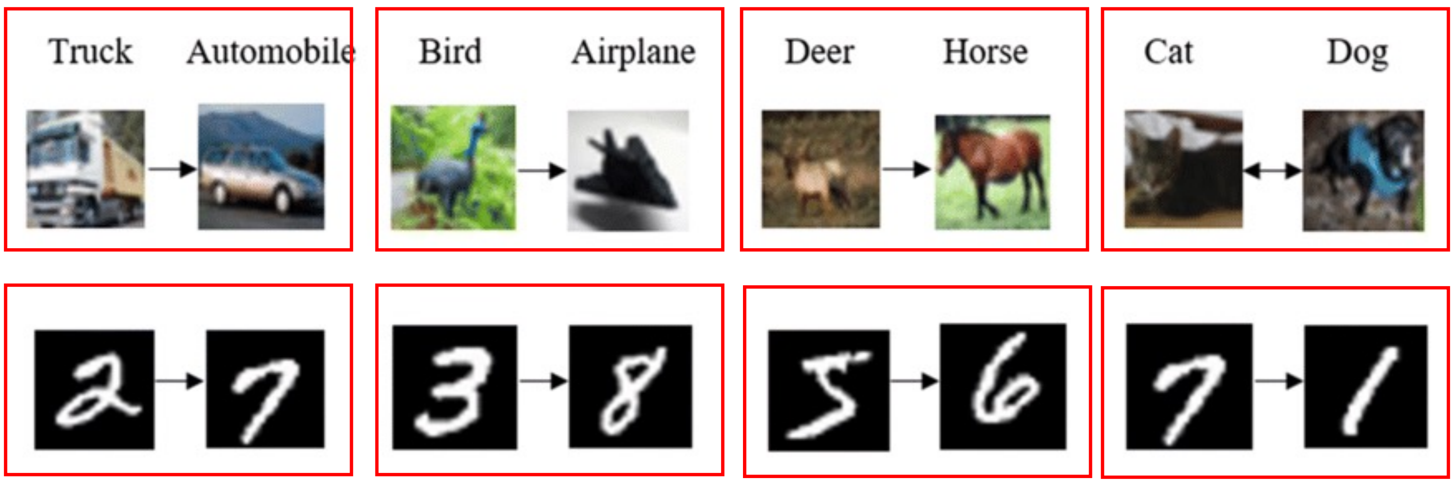
[Random Noise]

이미지출처: Ahmed, A., Yousif, H., & He, Z. (2021). Ensemble diversified learning for image classification with noisy labels. Multimedia Tools and Applications, 80, 20759-20772.

Noise Transition Matrix

❖ Noisy Label의 종류

- **Random Noise:** Instance Feature이나 Class에 영향을 받지 않음
  - 아무런 규칙없이 일정 확률로 잘못된 Label을 가지는 데이터
- **Instance-Independent Label Noise :** Instance Feature와는 독립적이지만 Class에 영향을 받음
  - 유사한 특정 Class 간의 인식 오인으로 발생한 Noisy Label



[Instance-Independent Label Noise]

이미지출처: Ahmed, A., Yousif, H., & He, Z. (2021). Ensemble diversified learning for image classificati
on with noisy labels. Multimedia Tools and Applications, 80, 20759-20772.

Noise Transition Matrix

Instance-Independent Label Noise의 존재로 인해 발생한 Class 간 오인 확률

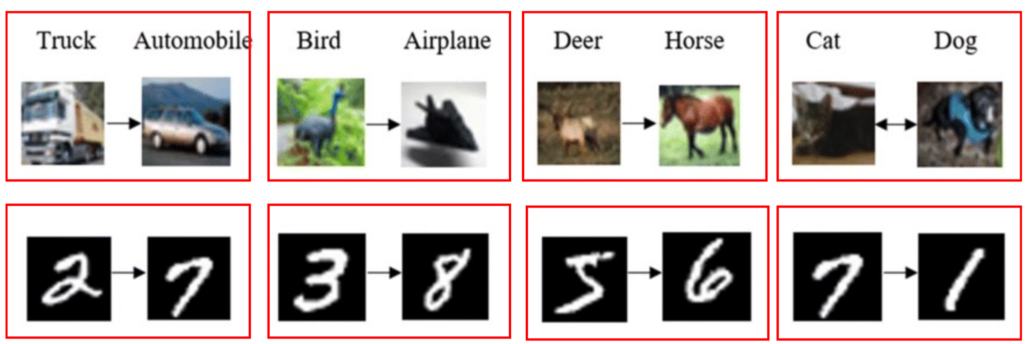
$$\underset{(k\times k)}{Transition\ Matrix}\quad T = \begin{pmatrix} T_{11} & T_{12} & \cdots & T_{1k} \\ T_{21} & T_{22} & \cdots & T_{2k} \\ \vdots & \vdots & & \vdots \\ T_{k1} & T_{k2} & \cdots & T_{kk} \end{pmatrix}$$

True Label

$$P\left(y^f = j \middle| \hat{y} = i\right) = T_{ij}$$

Predicted Label

실제 레이블 i를 레이블 j로 오인한 확률



[Instance-Independent Label Noise]

이미지 출처: Ahmed, A., Yousif, H., & He, Z. (2021). Ensemble diversified learning for image classification with noisy labels. Multimedia Tools and Applications, 80, 20759-20772.

Noise Transition Matrix

❖ Using Trusted Data to Train Deep Networks on Labels Corrupted by Severe Noise

- 2018년 NeurIPS 에 발표된 논문
- 2023년 06월 02일 기준 인용 횟수 : 463회

Using Trusted Data to Train Deep Networks on Labels Corrupted by Severe Noise

Dan Hendrycks*
University of California, Berkeley
hendrycks@berkeley.edu

Mantas Mazeika*
University of Chicago
mantas@ttic.edu

Duncan Wilson
Foundational Research Institute
duncanw@nevada.unr.edu

Kevin Gimpel
Toyota Technological Institute at Chicago
kgimpel@ttic.edu

Abstract

The growing importance of massive datasets used for deep learning makes robustness to label noise a critical property for classifiers to have. Sources of label noise include automatic labeling, non-expert labeling, and label corruption by data poisoning adversaries. Numerous previous works assume that no source of labels can be trusted. We relax this assumption and assume that a small subset of the training data is trusted. This enables substantial label corruption robustness performance gains. In addition, particularly severe label noise can be combated by using a set of trusted data with clean labels. We utilize trusted data by proposing a loss correction technique that utilizes trusted examples in a data-efficient manner to mitigate the effects of label noise on deep neural network classifiers. Across vision and natural language processing tasks, we experiment with various label noises at several strengths, and show that our method significantly outperforms existing methods.

Noise Transition Matrix

❖ Using Trusted Data to Train Deep Networks on Labels Corrupted by Severe Noise

$< Corruption\ probabilities >$

True Label

$$P(\tilde{y} = j | y = i) = C_{ij} = T_{ij}$$

Predicted Label

실제 레이블 i를 레이블 j로 오인한 확률

Noisy Label이 포함된 데이터로 학습한 모델

입력에 x에 맞는 True Label을 출력하는 모델

$$p(\tilde{y}|x) = \sum_{y=1}^{K} p(\tilde{y}|y,x)\, p(y|x)$$

C_{ij} 확률에 의한 Label Corruption 발생

위 관계의 반영을 통해 Noisy Label에 오염되지 않은

입력에 x에 맞는 True Label을 출력하는 모델 p(y|x)를 찾고자 함

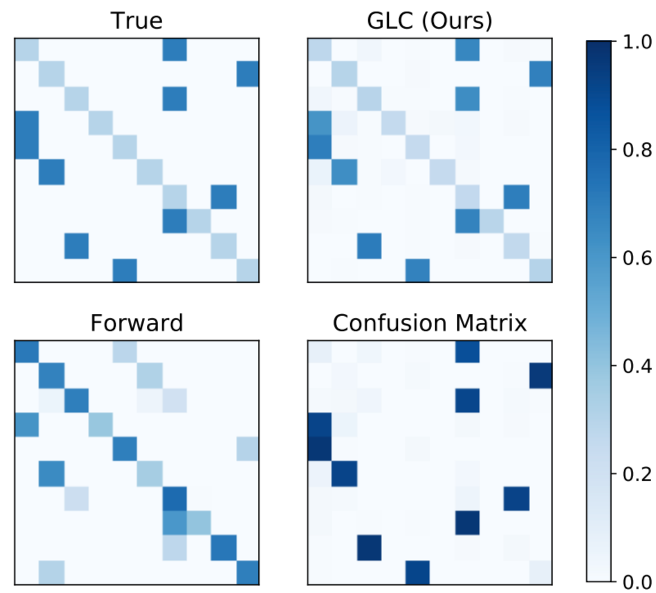


Figure 1: A label corruption matrix (top left) and three matrix estimates for a corrupted CIFAR-10 dataset. Entry C_{ij} is the probability that a label of class i is corrupted to class j, or symbolically $C_{ij} = p(\tilde{y} = j | y = i)$.

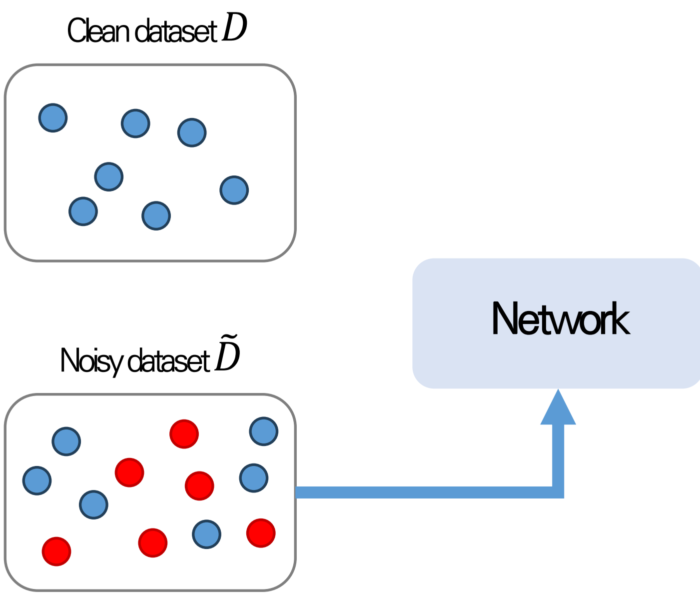
Noise Transition Matrix

❖ Using Trusted Data to Train Deep Networks on Labels Corrupted by Severe Noise

- Gold Loss Correction (GLC) Algorithm

Step 1	Noisy Label 데이터가 포함된 $\widetilde{\mathcal{D}}$으로 Network 학습

Transition Matrix 구하는 단계



Algorithm Gold Loss Correction (GLC)

Input: Trusted data $\mathcal{D}$, untrusted data $\widetilde{\mathcal{D}}$, loss ℓ
Train network $f(x) = \widehat{p}(\widetilde{y}|x;\theta) \in \mathbb{R}^K$ on $\widetilde{\mathcal{D}}$
Fill $\widehat{C} \in \mathbb{R}^{K\times K}$ with zeros
for $k = 1, \dots, K$ **do**
  `num_examples` $= 0$
  for $(x_i, y_i) \in \mathcal{D}$ such that $y_i = k$ **do**
    `num_examples` $\mathrel{+}= 1$
    $\widehat{C}_{k\bullet} \mathrel{+}= f(x_i)$ {add $f(x_i)$ to kth row}
  end for
  $\widehat{C}_{k\bullet} \mathrel{/}=$ `num_examples`
end for
Initialize new model $g(x) = \widehat{p}(y|x;\theta)$
Train with $\ell(g(x), y)$ on $\mathcal{D}$, $\ell(\widehat{C}^\top g(x), \widetilde{y})$ on $\widetilde{\mathcal{D}}$
Output: Model $\widehat{p}(y|x;\theta)$

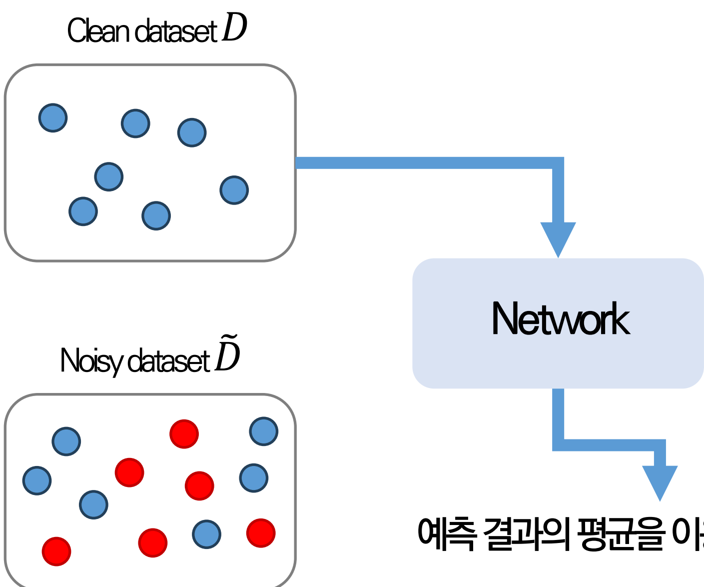
Noise Transition Matrix

❖ Using Trusted Data to Train Deep Networks on Labels Corrupted by Severe Noise

- Gold Loss Correction (GLC) Algorithm

Step 2	Clean dataset으로 Network 예측 결과 산출하여 C_{ij} 구축

Transition Matrix 구하는 단계



예측 결과의 평균을 이용하여 C_{ij} 구축

$$\begin{pmatrix} C_{11} & C_{12} & \cdots & C_{1k} \\ C_{21} & C_{22} & \cdots & C_{2k} \\ \vdots & \vdots & & \vdots \\ C_{k1} & C_{k2} & \cdots & C_{kk} \end{pmatrix}$$

Algorithm GOLD LOSS CORRECTION (GLC)

Input: Trusted data $\mathcal{D}$, untrusted data $\widetilde{\mathcal{D}}$, loss ℓ
Train network $f(x) = \widehat{p}(\widetilde{y}|x;\theta) \in \mathbb{R}^K$ on $\widetilde{\mathcal{D}}$
Fill $\widehat{C} \in \mathbb{R}^{K\times K}$ with zeros
for $k = 1, \ldots, K$ **do**
  `num_examples` $= 0$
  for $(x_i, y_i) \in \mathcal{D}$ such that $y_i = k$ **do**
    `num_examples` $+= 1$
    $\widehat{C}_{k\bullet}$ $+= f(x_i)$ {add $f(x_i)$ to kth row}
  end for
  $\widehat{C}_{k\bullet}$ $/=$ `num_examples`
end for
Initialize new model $g(x) = \widehat{p}(y|x;\theta)$
Train with $\ell(g(x), y)$ on $\mathcal{D}$, $\ell(\widehat{C}^\top g(x), \widetilde{y})$ on $\widetilde{\mathcal{D}}$
Output: Model $\widehat{p}(y|x;\theta)$

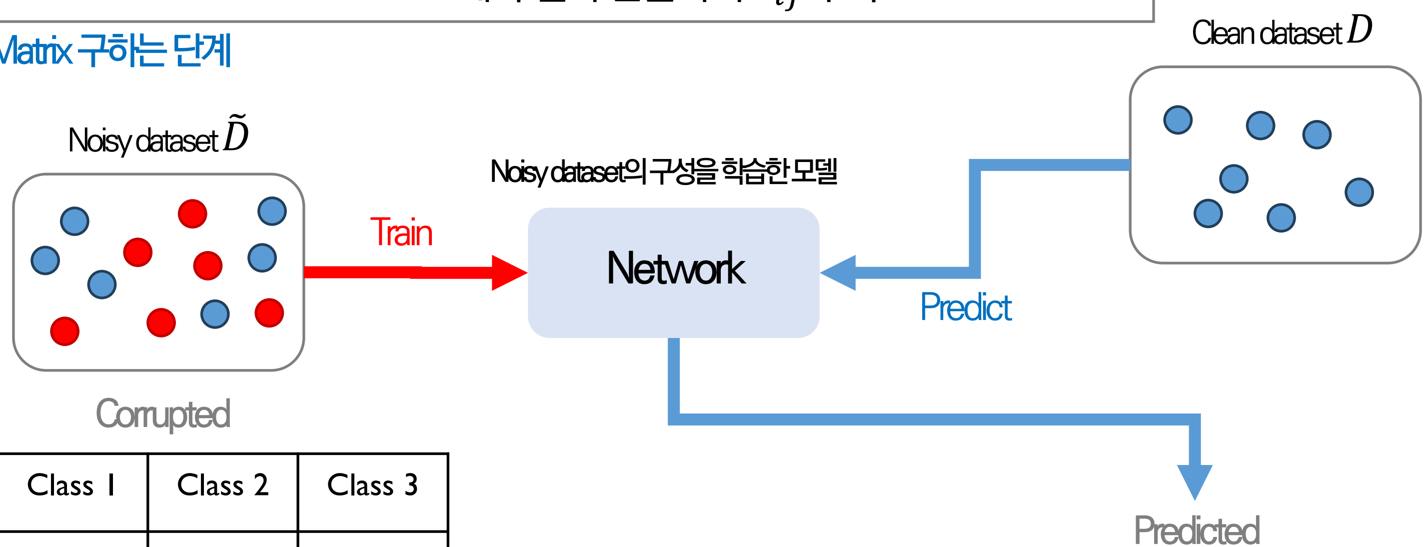
Noise Transition Matrix

❖ Using Trusted Data to Train Deep Networks on Labels Corrupted by Severe Noise

- Gold Loss Correction (GLC) Algorithm

Step 2	Clean dataset으로 Network 예측 결과 산출하여 C_{ij} 구축

Transition Matrix 구하는 단계



True

	Class 1	Class 2	Class 3
Class 1	60%	30%	10%
Class 2	20%	80%	0%
Class 3	10%	0%	90%

Noisy dataset의 구성

True

$$C_{ij} = \begin{pmatrix} 0.6 & 0.3 & 0.1 \\ 0.2 & 0.8 & 0.0 \\ 0.1 & 0.0 & 0.9 \end{pmatrix}$$

예측 결과의 평균

Noise Transition Matrix

❖ Using Trusted Data to Train Deep Networks on Labels Corrupted by Severe Noise

- Gold Loss Correction (GLC) Algorithm

Step 2	Clean dataset으로 Network 예측 결과 산출하여 C_{ij} 구축

Transition Matrix 구하는 단계

$$C_{ij} = \begin{pmatrix} 0.6 & 0.3 & 0.1 \\ 0.2 & 0.8 & 0.0 \\ 0.1 & 0.0 & 0.9 \end{pmatrix}$$

예측 결과의 평균

Transition Matrix 구축 완료

C_{11} : 실제 Class 1 의 데이터가 입력되었을 때, Noisy Dataset으로 오염된 모델이 Class 1로 예측할 확률
C_{12} : 실제 Class 1 의 데이터가 입력되었을 때, Noisy Dataset으로 오염된 모델이 Class 2로 예측할 확률
C_{13} : 실제 Class 1 의 데이터가 입력되었을 때, Noisy Dataset으로 오염된 모델이 Class 3으로 예측할 확률
C_{21} : 실제 Class 2의 데이터가 입력되었을 때, Noisy Dataset으로 오염된 모델이 Class 1로 예측할 확률
⋮
C_{33} : 실제 Class 3 의 데이터가 입력되었을 때, Noisy Dataset으로 오염된 모델이 Class 3으로 예측할 확률

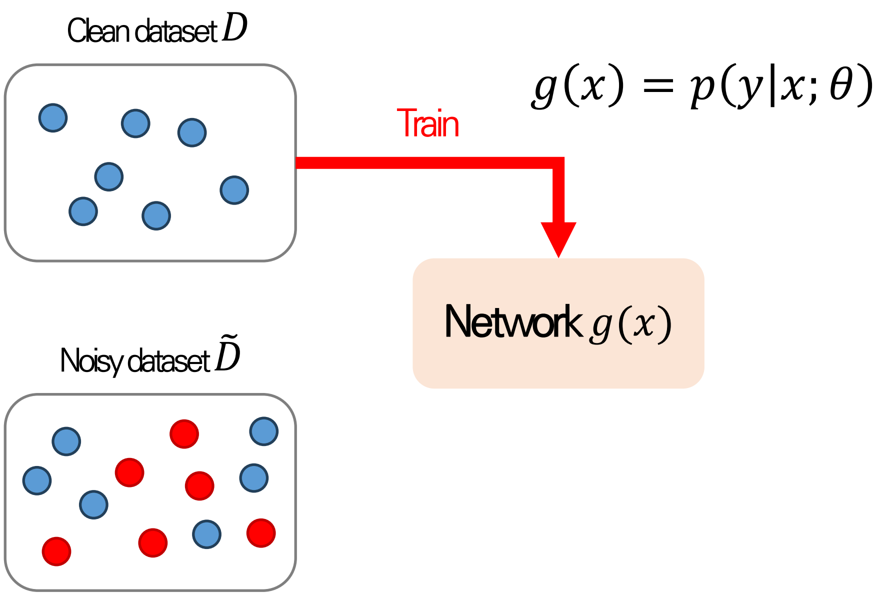
Noise Transition Matrix

❖ Using Trusted Data to Train Deep Networks on Labels Corrupted by Severe Noise

- Gold Loss Correction (GLC) Algorithm

Step 3	Clean dataset D를 이용하여 초기 모델 학습

최종모델 학습 단계



Algorithm GOLD LOSS CORRECTION (GLC)

Input: Trusted data $\mathcal{D}$, untrusted data $\widetilde{\mathcal{D}}$, loss ℓ
Train network $f(x) = \widehat{p}(\widetilde{y}|x;\theta) \in \mathbb{R}^K$ on $\widetilde{\mathcal{D}}$
Fill $\widehat{C} \in \mathbb{R}^{K\times K}$ with zeros
for $k = 1, \ldots, K$ **do**
    num_examples $= 0$
    for $(x_i, y_i) \in \mathcal{D}$ such that $y_i = k$ **do**
        num_examples $+= 1$
        $\widehat{C}_{k\bullet} += f(x_i)$ {add $f(x_i)$ to kth row}
    end for
    $\widehat{C}_{k\bullet}$ /= num_examples
end for
Initialize new model $g(x) = \widehat{p}(y|x;\theta)$
Train with $\ell(g(x), y)$ on $\mathcal{D}$, $\ell(\widehat{C}^{\mathsf{T}} g(x), \widetilde{y})$ on $\widetilde{\mathcal{D}}$
Output: Model $\widehat{p}(y|x;\theta)$

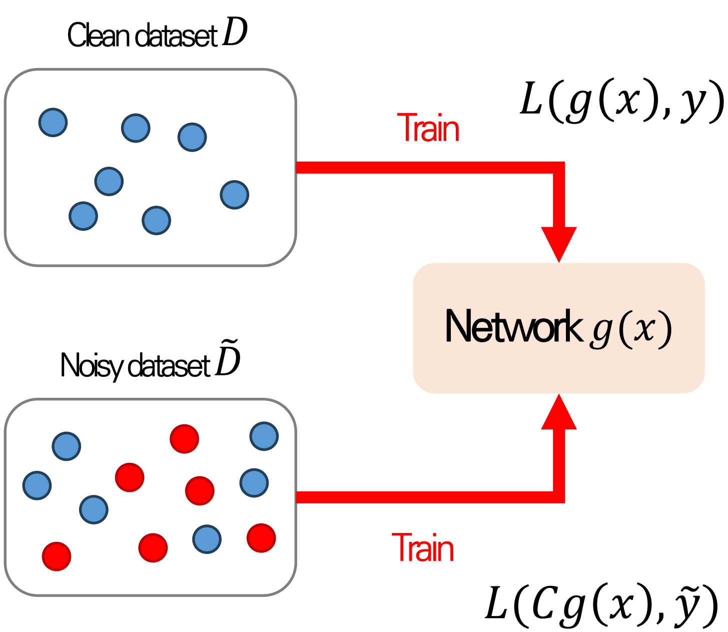
Noise Transition Matrix

❖ Using Trusted Data to Train Deep Networks on Labels Corrupted by Severe Noise

- Gold Loss Correction (GLC) Algorithm

Step 4	모든 dataset을 이용해 최종 모델 학습

최종모델 학습 단계



Algorithm GOLD LOSS CORRECTION (GLC)

Input: Trusted data $\mathcal{D}$, untrusted data $\widetilde{\mathcal{D}}$, loss ℓ
Train network $f(x) = \widehat{p}(\widetilde{y}|x;\theta) \in \mathbb{R}^K$ on $\widetilde{\mathcal{D}}$
Fill $\widehat{C} \in \mathbb{R}^{K \times K}$ with zeros
for $k = 1, \ldots, K$ **do**
    $\texttt{num\_examples} = 0$
    for $(x_i, y_i) \in \mathcal{D}$ such that $y_i = k$ **do**
        $\texttt{num\_examples} \mathrel{+}= 1$
        $\widehat{C}_{k\bullet} \mathrel{+}= f(x_i)$ {add $f(x_i)$ to kth row}
    end for
    $\widehat{C}_{k\bullet} \mathrel{/}= \texttt{num\_examples}$
end for
Initialize new model $g(x) = \widehat{p}(y|x;\theta)$
Train with $\ell(g(x), y)$ on $\mathcal{D}$, $\ell(\widehat{C}^{\top} g(x), \widetilde{y})$ on $\widetilde{\mathcal{D}}$
Output: Model $\widehat{p}(y|x;\theta)$

$$Loss\ function = L(g(x), y) + L(Cg(x), \tilde{y})$$

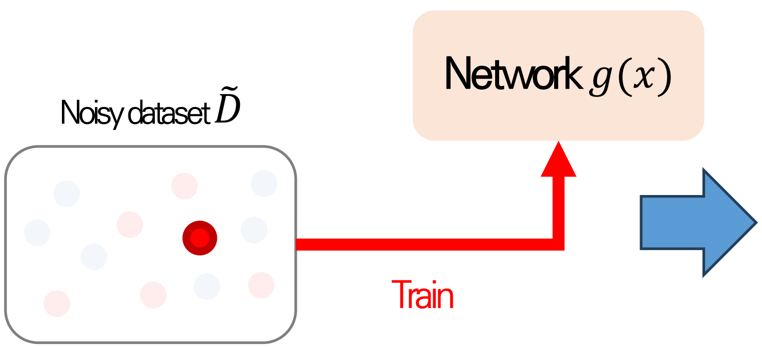
Noise Transition Matrix

❖ Using Trusted Data to Train Deep Networks on Labels Corrupted by Severe Noise

- Gold Loss Correction (GLC) Algorithm

Step 4	모든 dataset을 이용해 최종 모델 학습

최종모델 학습 단계



실제로는 Class 1이지만, Class 2로 Labeling된 샘플

$$g(x_k) = \begin{pmatrix} 0.8 \\ 0.2 \\ 0.0 \end{pmatrix} \qquad \tilde{y}_k = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}$$

$$\boldsymbol{C}g(x_k) = \begin{pmatrix} 0.6 & 0.3 & 0.1 \\ 0.2 & 0.8 & 0.0 \\ 0.1 & 0.0 & 0.9 \end{pmatrix} \begin{pmatrix} 0.8 \\ 0.2 \\ 0.0 \end{pmatrix} = \begin{pmatrix} 0.6\times0.8 + 0.3\times0.2 + 0.1\times0.0 \\ 0.2\times0.8 + 0.8\times0.2 + 0.0\times0.0 \\ 0.1\times0.8 + 0.0\times0.2 + 0.9\times0.0 \end{pmatrix}$$

$$= \begin{pmatrix} 0.54 \\ 0.32 \\ 0.08 \end{pmatrix}$$

Noisy sample에 대한 출력확률 smoothing

해당 Noisy sample에 대한 모델의 update 영향력을 낮추는 결과

Noisy dataset의 손상 가능성을 반영하여 Network가 올바른 방향으로 Noisy dataset을 학습할 수 있게 함

Noise Transition Matrix

❖ Training deep neural-networks using a noise adaptation layer.

- 2017년 ICLR에 발표된 논문
- 2023년 06월 02일 기준 인용 횟수 : **554회**

TRAINING DEEP NEURAL-NETWORKS USING A NOISE ADAPTATION LAYER

Jacob Goldberger & Ehud Ben-Reuven
Engineering Faculty, Bar-Ilan University,
Ramat-Gan 52900, Israel
jacob.goldberger@biu.ac.il,udi.benreuven@gmail.com

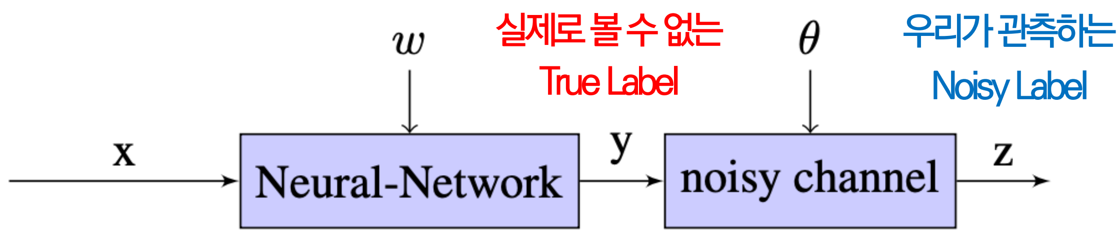
ABSTRACT

The availability of large datsets has enabled neural networks to achieve impressive recognition results. However, the presence of inaccurate class labels is known to deteriorate the performance of even the best classifiers in a broad range of classification problems. Noisy labels also tend to be more harmful than noisy attributes. When the observed label is noisy, we can view the correct label as a latent random variable and model the noise processes by a communication channel with unknown parameters. Thus we can apply the EM algorithm to find the parameters of both the network and the noise and estimate the correct label. In this study we present a neural-network approach that optimizes the same likelihood function as optimized by the EM algorithm. The noise is explicitly modeled by an additional softmax layer that connects the correct labels to the noisy ones. This scheme is then extended to the case where the noisy labels are dependent on the features in addition to the correct labels. Experimental results demonstrate that this approach outperforms previous methods.

Noise Transition Matrix

❖ Training deep neural-networks using a noise adaptation layer.

- 앞의 논문과 같이, True Label y를 Noisy Label z로 매핑되는 함수(noisy channel)를 찾아내는 것이 목표



정확한 Label을 예측하는 모델

$$L(w,\theta) = \sum_{t=1}^{n} \log\left(\sum_{i=1}^{k} p(z_t|y_t = i;\theta)\,p(y_t = i|x_t;w)\right) \qquad (2)$$

클래스별 Label이 오염되는 관계
Transition Matrix

Using Trusted Data to Train Deep Networks on Labels Corrupted by Severe Noise $L(Cg(x), \tilde{y})$

Noise Transition Matrix

❖ Training deep neural-networks using a noise adaptation layer.

- Soft-max layer를 이용하여 y가 클래스별로 오염되는 관계를 쉽게 구현

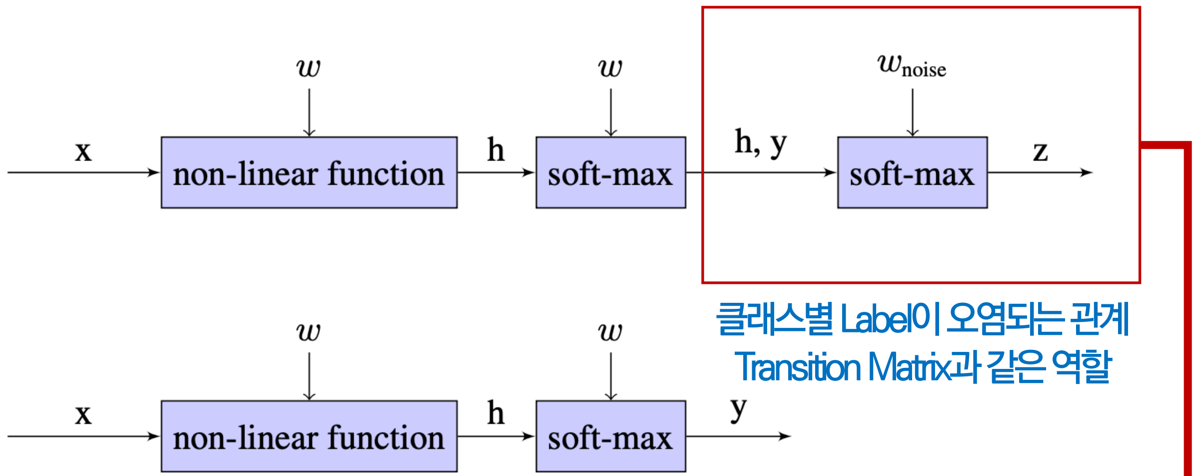


Figure 1: An illustration of the noisy-label neural network architecture for the training phase (above) and test phase (below).

$$p(z=j|y=i)=\theta(i,j)=\frac{\exp(b_{ij})}{\sum_l \exp(b_{il})}$$

앞선 논문과 동일하게 오염된 모델의 예측 결과(Confusion Matrix)를 통해 클래스 별로 y가 오염될 확률인 Transition Matrix $\boldsymbol{b_{ij}}$를 구함

GLC Algorithm의 과정을 softmax-layer로 구현하였다는 점을 제외하고는 GLC와 동일한 알고리즘

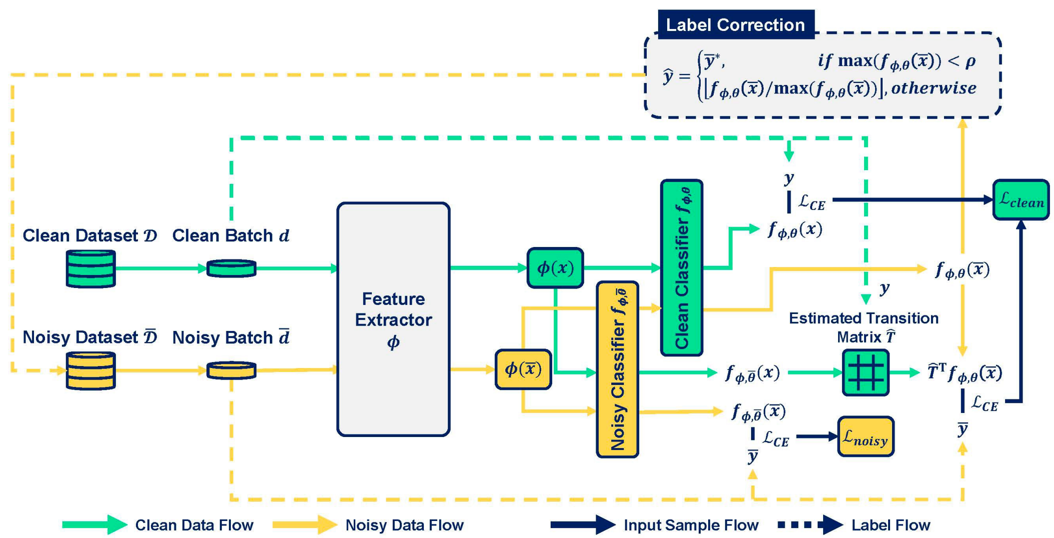
Noise Transition Matrix

❖ Noise Transition Matrix 관련 연구 방향

- Learning with Noisy Labels by Efficient Transition Matrix Estimation to Combat Label Miscorrection. In Computer Vision–ECCV 2022

$$\mathcal{L}_{clean} = \mathcal{L}_{CE}(f_{\phi,\theta}(x), y) + \boxed{\sum_{(\bar{x},\bar{y})\in\bar{d}} \left(\mathcal{L}_{CE}\left(\hat{T}^{\top} f_{\phi,\theta}(\bar{x}), \bar{y}\right)\right)}$$

Noisy Dataset에 대한 예측 결과에 Transition Matrix를 곱한 형태



Noise Transition Matrix

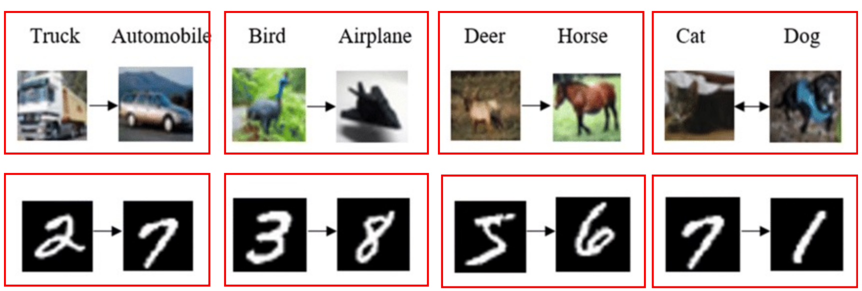
❖ Noise Transition Matrix 관련 연구 방향

- Li, Shikun, et al. "Estimating noise transition matrix with label correlations for noisy multi-label learning." Advances in Neural Information Processing Systems. 2022.
- Yao, Yu, et al. "Dual t: Reducing estimation error for transition matrix in label-noise learning." Advances in Neural Information Processing Systems. 2020.
- Zhu, Zhaowei, Jialu Wang, and Yang Liu. "Beyond images: Label noise transition ma trix estimation for tasks with lower-quality features." International Conference on M achine Learning. PMLR, 2022.
- Zhang, Yivan, Gang Niu, and Masashi Sugiyama. "Learning noise transition matrix fr om only noisy labels via total variation regularization." International Conference on Machine Learning. PMLR, 2021.

Robust Loss Regularization

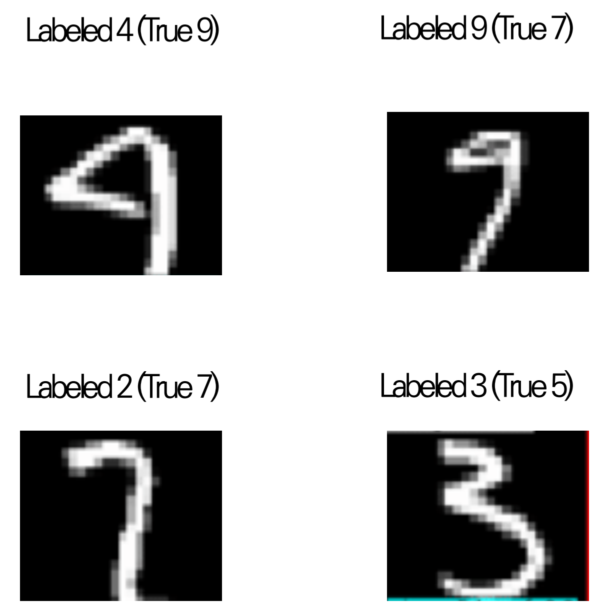
❖ Noisy Label의 종류

- **Instance-dependent Label Noise :** Instance Feature와 Class 모두에 영향을 받음
  - Class 간의 관계만으로는 설명이 되지 않는 Noisy Label도 포함



[Instance-Independent Label Noise]

이미지출처: Ahmed, A., Yousif, H., & He, Z. (2021). Ensemble diversified learning for image classification with noisy labels. Multimedia Tools and Applications, 80, 20759-20772.



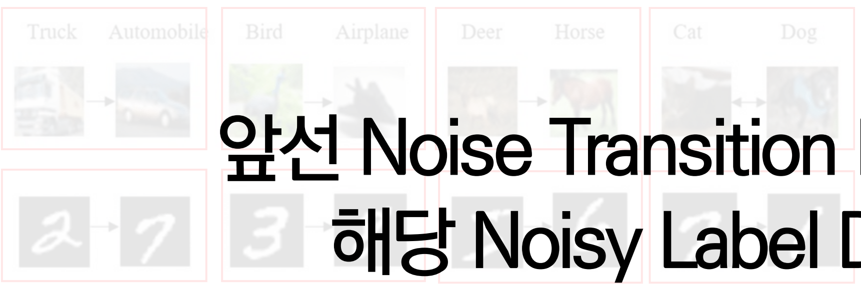
클래스 간의 관계만으로는 설명되지 않는 Noisy Label 케이스

실제 현실 데이터에는 클래스 간 오인률이 명확하지 않은 많은 종류의 Noisy Label이 있음

Robust Loss Regularization

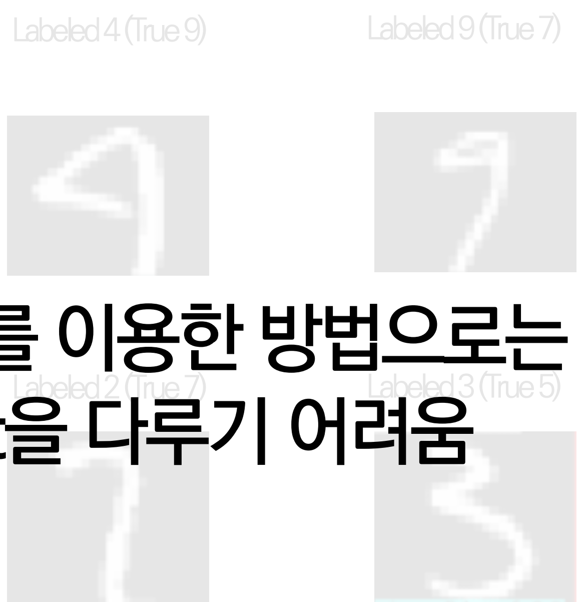
❖ Noisy Label의 종류

- Instance-dependent Label Noise : Instance Feature와 Class 모두에 영향을 받음
  - Class 간의 관계만으로는 설명이 되지 않는 Noisy Label도 포함



[Instance-Independent Label Noise]

이미지출처: Ahmed, A., Yousif, H., & He, Z. (2021). Ensemble diversified learning for image classification with noisy labels. Multimedia Tools and Applications, 80, 20759-20772.



클래스 간의 관계만으로는 설명되지 않는 Noisy Label 케이스

앞선 Noise Transition Matrix를 이용한 방법으로는 해당 Noisy Label Dataset을 다루기 어려움

실제 현실 데이터에는 클래스 간 오인률이 명확하지 않은 많은 종류의 Noisy Label이 있음

Robust Loss Regularization

❖ Early–learning regularization prevents memorization of noisy labels

- 2020년 NeurIPS에 발표된 논문
- 2023년 06월 02일 기준 인용 횟수 : **303회**

Early-Learning Regularization Prevents Memorization of Noisy Labels

Sheng Liu
Center for Data Science
New York University
shengliu@nyu.edu

Jonathan Niles-Weed
Center for Data Science, and
Courant Inst. of Mathematical Sciences
New York University
jnw@cims.nyu.edu

Narges Razavian
Department of Population Health, and
Department of Radiology
NYU School of Medicine
narges.razavian@nyulangone.org

Carlos Fernandez-Granda
Center for Data Science, and
Courant Inst. of Mathematical Sciences
New York University
cfgranda@cims.nyu.edu

Abstract

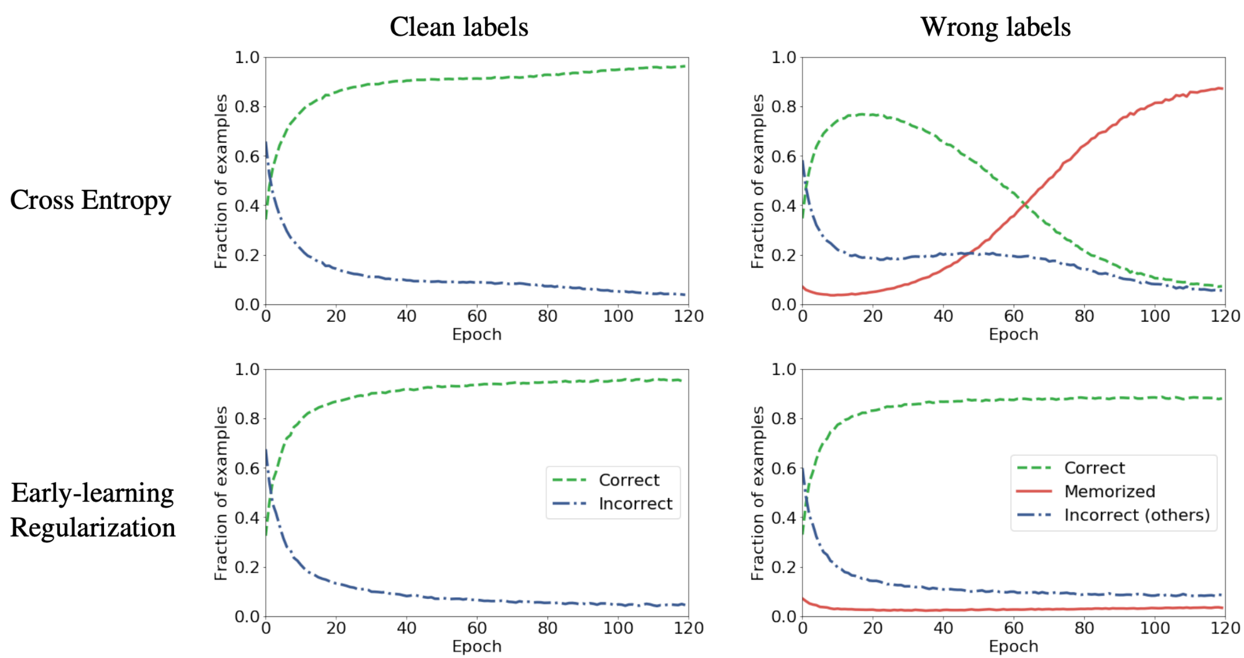
We propose a novel framework to perform classification via deep learning in the presence of noisy annotations. When trained on noisy labels, deep neural networks have been observed to first fit the training data with clean labels during an "early learning" phase, before eventually memorizing the examples with false labels. We prove that early learning and memorization are fundamental phenomena in high-dimensional classification tasks, even in simple linear models, and give a theoretical explanation in this setting. Motivated by these findings, we develop a new technique for noisy classification tasks, which exploits the progress of the early learning phase. In contrast with existing approaches, which use the model output during early learning to detect the examples with clean labels, and either ignore or attempt to correct the false labels, we take a different route and instead capitalize on early learning via regularization. There are two key elements to our approach. First, we leverage semi-supervised learning techniques to produce target probabilities based on the model outputs. Second, we design a regularization term that steers the model towards these targets, implicitly preventing memorization of the false labels. The resulting framework is shown to provide robustness to noisy annotations on several standard benchmarks and real-world datasets, where it achieves results comparable to the state of the art.

Robust Loss Regularization

❖ Early-learning regularization prevents memorization of noisy labels

- Clean dataset의 경우, 초기에 모델이 빠르게 학습하여 gradient가 Noisy Dataset에 비해 크기가 작아짐
- 이에 따라 Noisy Dataset의 gradient 크기가 우세해지고 모델은 점차 Noisy Dataset을 기억하는 방향으로 학습

모델이 학습 단계를 거치면서, Noisy Label data를 기억(overfitting)하는 것을 방지해야 함



Correct(초록선) : 모델이 실제 정답을 올바르게 예측
Incorrect(파란선) : 모델이 실제 정답을 틀리게 예측
Memorized(빨간선) : 모델의 예측 값이 잘못된 레이블과 같을 경우

Robust Loss Regularization

❖ Early-learning regularization prevents memorization of noisy labels

- ELR(Early-learning Regularization)을 통한 Memorization 방지

Regularization Term

$$\mathcal{L}_{\text{ELR}}(\Theta) := \mathcal{L}_{\text{CE}}(\Theta) + \boxed{\frac{\lambda}{n}\sum_{i=1}^{n}\log\left(1 - \langle \mathbf{p}^{[i]}, \mathbf{t}^{[i]} \rangle\right)} .$$

$p^{[i]}$: 모델 학습 중 모델이 입력 x_i에 대한 예측 확률값

$t^{[i]}$: 학습 과정 동안 $p^{[i]}$ 의 값들을 평균해서 산출

간단한 구현과 빠른 계산의 장점 : Training Set에 대해 $t^{[i]}$ 값만 저장하고 있으면 됨.

Robust Loss Regularization

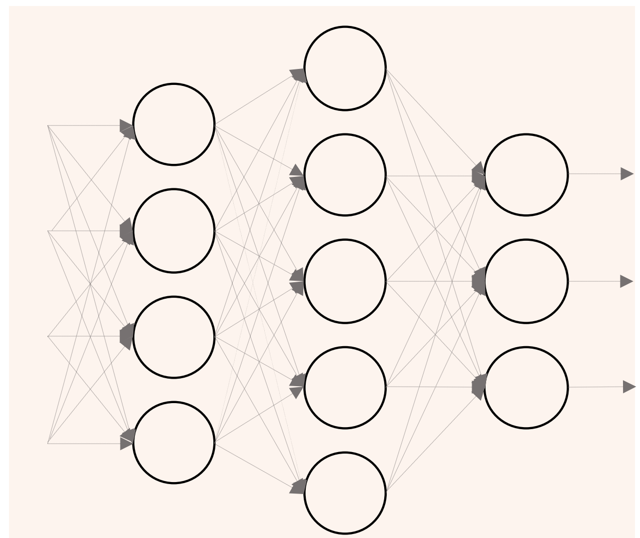
❖ Early–learning regularization prevents memorization of noisy labels

- ELR(Early–learning Regularization) 작동 원리

$$\mathcal{L}_{\mathrm{ELR}}(\Theta) := \mathcal{L}_{\mathrm{CE}}(\Theta) + \frac{\lambda}{n}\sum_{i=1}^{n}\log\left(1 - \langle \mathbf{p}^{[i]}, \mathbf{t}^{[i]} \rangle\right).$$

Input X

…



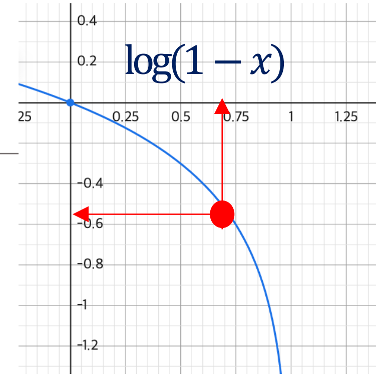
Neural Net

Output $\boldsymbol{Y}$

$\boldsymbol{Y}$
강아지
토끼
강아지
…
…
토끼
고양이
고양이

강아지,토끼,고양이를 분류하는 모델을 가정

Robust Loss Regularization

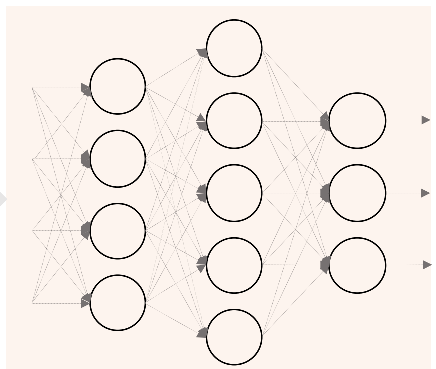


❖ Early–learning regularization prevents memorization of noisy labels

- ELR(Early–learning Regularization) 작동 원리

$$\mathcal{L}_{\mathrm{ELR}}(\Theta) := \mathcal{L}_{\mathrm{CE}}(\Theta) + \frac{\lambda}{n}\sum_{i=1}^{n}\log\left(1-\langle \mathbf{p}^{[i]}, \mathbf{t}^{[i]}\rangle\right).$$

Sample x_i



$$p^{[i]} = \begin{bmatrix} 0.8 \\ 0.2 \\ 0.0 \end{bmatrix}$$

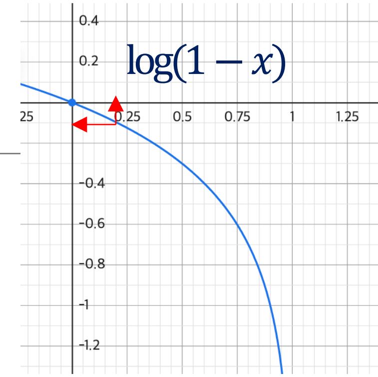
강아지로 예측할 확률
고양이로 예측할 확률
토끼로 예측할 확률

$$t^{[i]} = \begin{bmatrix} 0.9 \\ 0.1 \\ 0.0 \end{bmatrix}$$

$$\langle p^{[i]}, t^{[i]} \rangle = 0.74$$

한 sample x_i에 대해 모델 학습 과정 중 예측 확률값 $p^{[i]}$ 및 지난 학습 과정 동안의 평균 예측 확률값 $t^{[i]}$을 산출

Robust Loss Regularization



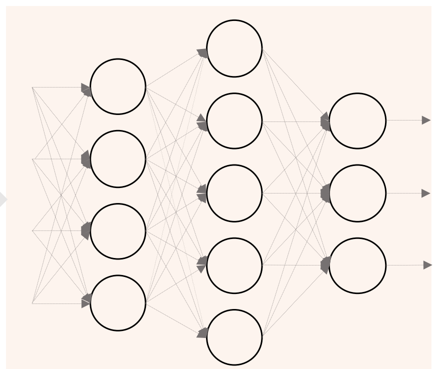
❖ Early–learning regularization prevents memorization of noisy labels

- ELR(Early–learning Regularization) 작동 원리

$$\mathcal{L}_{\mathrm{ELR}}(\Theta) := \mathcal{L}_{\mathrm{CE}}(\Theta) + \frac{\lambda}{n}\sum_{i=1}^{n}\log\left(1-\langle \mathbf{p}^{[i]}, \mathbf{t}^{[i]}\rangle\right).$$

$$p^{[i]} = \begin{bmatrix}0.8\\0.2\\0.0\end{bmatrix}$$

Sample x_i



$$p^{[i]} = \begin{bmatrix}0.1\\0.8\\0.1\end{bmatrix}$$

강아지로 예측할 확률

고양이로 예측할 확률

토끼로 예측할 확률

$$\langle p^{[i]}, t^{[i]}\rangle = 0.74$$

$$t^{[i]} = \begin{bmatrix}0.9\\0.1\\0.0\end{bmatrix} \quad \langle p^{[i]}, t^{[i]}\rangle = 0.17$$

현재 예측값 $p^{[i]}$ 이 기존 예측값 $t^{[i]}$와 크게 달라질 경우,

Regularization Term에 의해 전체 Loss Function의 크게 높임

Robust Loss Regularization

❖ Early-learning regularization prevents memorization of noisy labels

- Loss Function 변형 방법론이므로, 기존 Co-learning, Mix-up, weight averaging 등의 방법 조합 가능
- 본 논문에서는 지난 세미나에서 다룬었던 Co-learning, Mix-up의 방법론을 조합하여 ELR+로 칭함

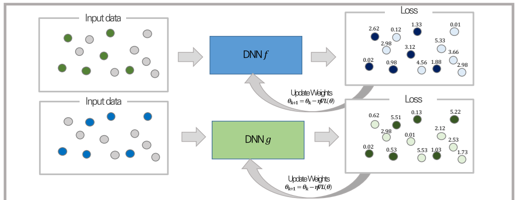
지난 세미나 :
http://dmqa.korea.ac.kr/activity/seminar/377

ELR

$$\mathcal{L}_{\mathrm{ELR}}(\Theta) := \mathcal{L}_{\mathrm{CE}}(\Theta) + \frac{\lambda}{n}\sum_{i=1}^{n}\log\left(1 - \langle \mathbf{p}^{[i]}, \mathbf{t}^{[i]} \rangle\right)$$

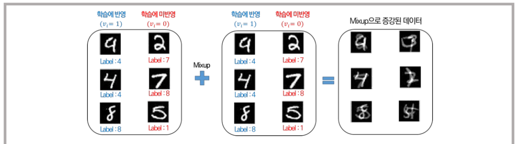
+

〈 Co-learning 〉



3. 서로 다른 네트워크에서 선별된 small-loss instance들로 모델 업데이트

〈 Mix-up 〉



Mixup을 통해 학습에 반영된 Noisy Label 데이터의 영향력을 줄이고,
학습 샘플 수 증강을 통해 모델 일반화 성능을 확보할 수 있다.

=

ELR+

Robust Loss Regularization

❖ Early-learning regularization prevents memorization of noisy labels

- 대표적인 Noisy Label Benchmark Dataset에서 ELR+가 기존 SOTA 대비 우수한 성능을 내는 것을 입증함

실험 데이터 : CIFAR-10N & CIFAR-100N

			Cross entropy	Co-teaching+ [52]	Mixup [55]	PENCIL [51]	MD-DYR-SH [2]	DivideMix [22]	ELR+	ELR+*
CIFAR-10	Sym. label noise	20%	86.8	89.5	95.6	92.4	94.0	**96.1**	94.6	95.8
		50%	79.4	85.7	87.1	89.1	92.0	94.6	93.8	**94.8**
		80%	62.9	67.4	71.6	77.5	86.8	93.2	91.1	**93.3**
		90%	42.7	47.9	52.2	58.9	69.1	76.0	75.2	**78.7**
	Asym.	40%	83.2	-	-	88.5	87.4	**93.4**	92.7	93.0
CIFAR-100	Sym. label noise	20%	62.0	65.6	67.8	69.4	73.9	77.3	77.5	**77.6**
		50%	46.7	51.8	57.3	57.5	66.1	**74.6**	72.4	73.6
		80%	19.9	27.9	30.8	31.1	48.2	60.2	58.2	**60.8**
		90%	10.1	13.7	14.6	15.3	24.3	31.5	30.8	**33.4**
	Asym.	40%	-	-	-	-	-	72.1	76.5	**77.5**

Table 2: Comparison with state-of-the-art methods on CIFAR-10 and CIFAR-100 with symmetric and asymmetric noise. For ELR+, we use 10% of the training set for validation, and treat the validation set as a held-out test set. The result for DivideMix on CIFAR-100 with 40% asymmetric noise was obtained using publicly available code. The rest of the results are taken from [22], which reports the highest accuracy observed on the validation set during training. We also report the performance of ELR+ under this metric on the rightmost column (ELR+*).

Robust Loss Regularization

❖ Early-learning regularization prevents memorization of noisy labels

- 대표적인 Noisy Label Benchmark Dataset에서 ELR+가 기존 SOTA 대비 우수한 성능을 내는 것을 입증함

실험 데이터 : Clothing 1M

CE	Forward [31]	GCE [56]	SL [45]	Joint-Optim [38]	DivideMix [22]	ELR	ELR+
69.10	69.84	69.75	71.02	72.16	74.76	72.87	**74.81**

Table 3: Comparison with state-of-the-art methods in test accuracy (%) on Clothing1M. All methods use a ResNet-50 architecture pretrained on ImageNet. Results of other methods are taken from the original papers (except for GCE, which is taken from [45]).

Conclusion

연구 동향 및 결론

❖ 다양한 접근 방식으로 Noisy Label Problem이 연구되고 있음

- Loss value Adjustment (2022.08.28 DMQA Open-Seminar)
- Mixup (2022.08.28 DMQA Open-Seminar)
- Co-learning (2022.08.28 DMQA Open-Seminar)
- Noise Transition Matrix
- Robust Loss Regularization

Thank you

본 세미나 내용에 대한 문의 사항이 있으시면
아래의 이메일 주소로 연락주시길 바랍니다.

E-mail : dawonksh@korea.ac.kr

참고 문헌

- ✓ Hendrycks, D., Mazeika, M., Wilson, D., & Gimpel, K. (2018). Using trusted data to train deep networks on labels corrupted by severe noise. Advances in neural information processing systems, 31.
- ✓ Goldberger, J., & Ben-Reuven, E. (2017, April). Training deep neural-networks using a noise adaptation layer. In International conference on learning representations
- ✓ ELR : Liu, S., Niles-Weed, J., Razavian, N., & Fernandez-Granda, C. (2020). Early-learning re gularization prevents memorization of noisy labels. Advances in neural information processing systems, 33, 20331-20342.
- ✓ Chen, P., Liao, B. B., Chen, G., & Zhang, S. (2019, May). Understanding and utilizing deep ne ural networks trained with noisy labels. In International Conference on Machine Learning (pp. 1062-1070). PMLR.
- ✓ Kye, S. M., Choi, K., Yi, J., & Chang, B. (2022, October). Learning with Noisy Labels by Efficie nt Transition Matrix Estimation to Combat Label Miscorrection. In Computer Vision–ECCV 202 2: 17th European Conference, Tel Aviv, Israel, October 23–27, 2022, Proceedings, Part XXV (p p. 717-738). Cham: Springer Nature Switzerland.