IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Received 1 March 2023, accepted 4 April 2023, date of publication 24 April 2023, date of current version 2 May 2023.

*Digital Object Identifier 10.1109/ACCESS.2023.3269792*

RESEARCH ARTICLE

# Glomerular Lesion Recognition Based on Pathology Images With Annotation Noise via Noisy Label Learning

**JING LI[1,2,*], QIMING HE[3,*], YIQING LIU[3], YANXIA WANG[1,2], TIAN GUAN[3], JING YE[1,2], YONGHONG HE[3], AND ZHE WANG[1,2]**

[1]State Key Laboratory of Cancer Biology, Department of Pathology, Xijing Hospital, Fourth Military Medical University, Xi'an 710032, China
[2]School of Basic Medicine, Fourth Military Medical University, Xi'an 710032, China
[3]Institute of Biopharmaceutical and Health Engineering, Tsinghua Shenzhen International Graduate School, Shenzhen 518055, China

Corresponding authors: Yonghong He (heyh@sz.tsinghua.edu.cn) and Zhe Wang (zhwang@fmmu.edu.cn)

This work was supported by the National Science Foundation of China under Grant 61875102, in part by the National Science Foundation of China under Grant 61975089, in part by the Science and Technology Research Program of Shenzhen City under Grant JCYJ20200109110606054, and in part by the Science and Technology Research Program of Shenzhen City under Grant WDZC20200821141349001.

*Jing Li and Qiming He contributed equally to this work.

**ABSTRACT** Background: Glomerular lesion recognition is one of the most crucial steps in the diagnosis of kidney disease. Deep learning, which relies on large numbers of pathology images, assists pathologists to access glomerular lesions more efficiently, objectively and accurately. However, due to different pathological development of glomeruli, complicated lesion patterns, and limited resolution of pathology images, there is annotation noise in datasets, making the deep learning model under- or over-fit. Methods: In this paper, we propose a novel noisy label learning model for lesion recognition in glomerular datasets with annotation noise. The model integrates uncertainty-based noisy label discriminator, contrastive learning, and consistency regularization to achieve high signal-to-noise supervision, pathology feature extraction, and utilization of pathology images. Results: We constructed large-scale glomerular datasets from 870 kidney disease cases using different stainings including Periodic acid-Schiff (PAS), Masson Trichrome (MT) and Periodic Schiff-Methenamine (PASM). Intensive experiments demonstrated the superiority of the proposed model for glomerular lesion recognition compared to other methods, as 25% of the lesions had $f_1 - score$ above 85%, 43.75% had $f_1 - score$ above 80%, and 75% had $f_1 - score$ at or above 70%. Additionally, further experiments demonstrate the effectiveness of each module. Conclusions: The noisy label learning model proposed is able to recognize the most glomerular lesions, with the annotation noise discrimination and large amounts of pathology images utilization, laying the foundation for the development of computer-aided evaluation system for the renal pathology.

**INDEX TERMS** Glomerular Lesions recognition, pathology images, annotation noise, noisy label learning, deep learning.

## I. INTRODUCTION

According to statistics, about 10% of adults have chronic kidney disease (CKD), which severely threatens life and health [1]. The glomerulus is the basic unit of kidney and plays an essential role in reflecting the onset, development and progression of kidney disease. Pathologists need to recognize glomerular lesions using several stainings, including PAS, MT, PASM and Hematoxylin and Eosin (H&E). For a case, pathologists perform a comprehensive and objective analysis of dozens of glomeruli on tissues, which involves complex morphological aspects due to different stainings and the superposition of multiple lesions. This routine work could be more efficient and reproducible with computer assistance.

Commonly, labeling glomerular lesions is the first step of deep learning. Though annotation quality is crucial for

The associate editor coordinating the review of this manuscript and approving it for publication was Yongjie Li.

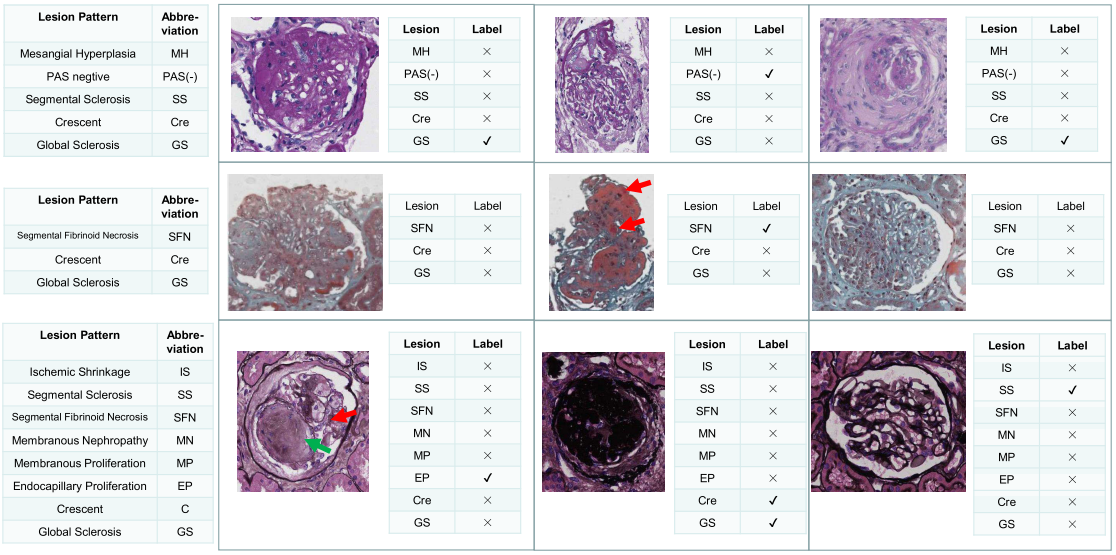



**FIGURE 1.** Examples of annotation noise. The lesion patterns are listed on the left. It is noted that the PAS(-) means that the mesangial matrix is significantly increased, with light staining under PAS staining. Rows 1-3 are from PAS, MT and PASM. ✓and × indicates the presence or absence of a lesion. Annotation noise arises from different causes. For example, in column 1 of row 1, glomerulus incorrectly labeled "GS" has developing glomerular change, as aproximately global sclerosing with some opening capillary lumens still being seen. In column 2 of row 2, due to the limited resolution, we cannot determine whether the red stained area is only "wire-loop deposits" or accompanied by SFN (red arrow). In column 1 of row 3, the glomerulus is labeled EP, but the ground truth is nodular sclerosis (green arrow) and a few foaming cells in the capillary (red arrow). This may be due to unobtrusive anomalous structures caused by the superposition of multiple lesions.

performance of computer assistance, noisy labels are present in the glomerulus dataset as shown in Fig. 1. First, some glomeruli are in the process of pathological development, whose lesions are atypical. Second, multiple lesions can coexist within a single glomerulus. Third, some fine structural alterations, such as vacuolations or spikes of basement membrane, could not be well recognized owing to limited resolution of light microscopy images. Comprehensive consideration of information from electron microscopy or immunofluorescence may be necessary in assessing the lesion. In summary, due to the above causes, annotation noise is generated to severely affect deep learning.

The current neglect of the significance of annotation noises in the datasets bottlenecks the performance of learning-based classification. When using a traditional loss function, the incorrect penalty to the deep model due to annotation noise can have a severe impact on model training. Several studies have proposed noise-robust loss [2], [3], [4], thus overcoming the effect of noise by robust penalty. However, such methods do not explicitly deal with the negative of noisy labels, and thus do not alleviate the problem of underfitting or overfitting when the dataset contains annotation noise.

In order to separate the noise from the dataset and thus explicitly perform direct processing of noisy samples, methods based on noise label recognition are proposed [5], [6]. After noisy data is collected, some work directly discards noisy samples and uses only clean labels for supervision [7]. This paradigm has the potential to discard both valuable correct labels and samples that help the network learn discriminative features. This brute-force approach can limit the generalization performance of the network. Other works have used label correction to correct the possible mislabeling of these samples. The label correction can generate pseudo-labels or label distributions of noisy samples, thus replacing the original potentially noisy labels [8], [9]. Although these methods can provide the network with supervised training with higher signal-to-noise ratio, incorrect label correction puts a constraint on the further improvement of the network performance.

To explicitly handle noisy labels and efficiently mining the glomerular images to improve the lesion recognition capability of the model, we propose a noisy label learning framework consisting of three modules. The noise label discriminator (NLD) robustly divides the dataset into a clean part and a noisy part based on sample uncertainty. NLD excludes the latter from the computation of the supervised classification loss, thereby avoiding the effect of annotation noise. The contrast learning (CL) module employs a Siamese network to extract sample features and constructs constraints on the model using the similarity relationship between samples. CL promotes the network to encode pathological features by forcing samples with higher similarity to be closer together in the feature space, and samples with lower similarity to be mutually exclusive. Compared to label correction, the unsupervised paradigm of CL can ensure the confidence of dependency information for network training to avoid semantic ambiguity. The consistency regularization (CR) module utilizes the teacher-student model to further enhance the network to capture lesion-related pathological features.

The contributions of this paper are summarized as follows:

- This paper presents a noisy label learning framework for accurate, generalizable glomerular lesion recognition

for PAS, MT, and PASM staining based on a large glomerular dataset with annotation noise.
- This paper designs an uncertainty-aware noisy label discriminator to achieve more efficient label discrimination and improve the signal-to-noise ratio for classification tasks.
- This paper uses Siamese model-based contrastive learning and teacher-student model-based consistency regularization to encode pathological features more efficiently and robustly.

The remainder of this paper is organized as follows. Section II reviews recent work on glomerular lesion recognition and noisy label learning with contrastive learning and sumesupervised learning. The details of the proposed method are described in Section III. The experimental results are provided in Section IV. Section V presents the discussion about our work. Finally, Sections VI concludes this paper.

## II. RELATED WORKS

### A. GLOMERULAR LESION RECOGNITION

Glomerulopathy is a manifestation of kidney disease at the nephron-unit level, and its assessment is necessary for a full diagnosis of nephrosis. Pathological image-based lesion recognition is one of the main tools. However, for one case, pathologists generally need to identify hundreds of glomeruli in detail, which involves complex morphological aspects of the glomerular lesion caused by different stainings and the superposition of multiple lesions. The present learning-based work is performed by full supervision on the construction of high-accuracy datasets. Some work is carried out for the glomerulus with a single lesion [10], [11]. [12] proposes to use uncertainty-aware module to improve the model's ability to recognize lesions. In other works, multi-stage method is used to construct the recognition of multiple lesions [13]. The processing of annotation noise in large-scale datasets needs to be further considered, especially in glomerular pathological images with extremely complex pathological patterns.

### B. NOISY LABEL LEARNING

Deep learning has been able to match human performance in relatively clean datasets [14], [15], [16], [17], as well as in computational pathology [18], [19]. However, for the histopathological image data, the heterogeneity of tissues and diseases leads to the difficulty, burden and subjectivity of annotation, which leads to the elevated cost of achieving high-quality annotation. Noisy label learning can help to address this issue, which can be divided into implicit and explicit handling of annotation noise. The implicit method is mainly to design noise robust loss function, so as to alleviate the non-robustness caused by only using traditional cross entropy loss to understand noisy labels [2], [3], [4]. Explicit methods will distinguish between clean and noisy labels and use different methods to process [20], [21], [22]. In previous studies, samples with noisy labels were directly discarded [7], which may lead to the loss of some valuable discriminant features during training, because such samples are more likely to be incorrectly labeled. In order to solve this problem, the label correction method was introduced [9], [23], which improved the efficiency of data mining in the network.

### C. CONTRASTIVE LEARNING

Contrastive learning is a paradigm of self-supervised learning that enables models to achieve similar performance to supervised learning [24], [25], [26]. The principle of contrastive learning is to narrow the distance between anchors and positive samples by designing a contrastive loss function and repel anchors and negative samples in the feature space. Without labels, contrastive learning can enhance the network to encode salient features and learn better feature representations.

### D. SEMI-SUPERVISED LEARNING

The proposal of semi-supervised learning is for the case where some samples are missing labels. In addition to supervised information from labeled samples, semi-supervised learning also uses all labeled and unlabeled samples to provide the network with additional knowledge about the data distribution, which can better help the network estimate decision boundaries. Semi-supervised methods include consistency regularization [27], self-training [28], collaborative training [29].

## III. METHODS

### A. DATASET

We collect kidney specimens of 870 patients at Xijing Hospital in Xi'an, China. The tissues were from the main nephropathy case, including DN (Diabetic Nephropathy), FSGS (Focal Segmental Glomerulosclerosis), AAGN (ANCA-Associated Glomerulonephritis), MN(Membranous Glomerulopathy), ORGN (Obesity-Related Glomerulopathy), AGBM (Anti-Glomerular Basement Membrane disease), LN (Lupus Nephritis), EPGN (Endocapillary Proliferative Glomerulonephritis), MPGN (Membranoproliferative Glomerulonephritis), TIN(Tubulointerstitial nephritis), CrGN (Crescentic Glomerulonephritis), IgAN (IgA Nephropathy) and HSP (Henoch-Schonlein Purpura). Each case was sliced on 12 consecutive levels, with both H&E and PAS having four levels and both MT and PASM having two. The tissues stained by PAS, MT, and PASM were scanned to obtain whole slide images (WSIs) by slide scanning image system of Shenzhen Shengqiang Technology Co.,. The dataset includes 842 PAS WSIs, 803 MT WSIs and 838 PASM WSIs.

We detected glomeruli at 5× equivalent magnification (1.68 um/pixel) using pretrained Mask R-CNN [30], and extracted glomerulus onto images at 20×. It is noted that the images were not intentionally selected to have complications such as large size differences and multiple lesions superimposed, and only one level for each staning is used. We extracted 17901, 15820 and 19197 glomeruli from PAS, MT and PASM, respectively, and randomly split the data

**TABLE 1.** The information of three datasets.

| PAS | MH | PAS(-) | SS | Cre | GS | NOA | Total |
|---|---|---|---|---|---|---|---|
| Train | 2681 | 106 | 80 | 171 | 1050 | 8195 | 12283 |
| Val | 674 | 29 | 22 | 31 | 241 | 2052 | 3049 |
| Test | 877 | 57 | 72 | 108 | 304 | 1151 | 2569 |

| MT | SFN | Cre | GS | NOA | Total |
|---|---|---|---|---|---|
| Train | 355 | 37 | 497 | 10345 | 11234 |
| Val | 102 | 10 | 125 | 2573 | 2810 |
| Test | 62 | 36 | 123 | 1555 | 1776 |

| PASM | IS | SS | SFN | MN | MP | EP | Cre | GS | NOA | Total |
|---|---|---|---|---|---|---|---|---|---|---|
| Train | 45 | 58 | 63 | 1648 | 421 | 66 | 242 | 1213 | 9535 | 13291 |
| Val | 21 | 17 | 16 | 404 | 98 | 18 | 51 | 291 | 2367 | 3283 |
| Test | 29 | 38 | 39 | 224 | 136 | 64 | 117 | 306 | 1670 | 2623 |

into training, validation and test sets according to the cases, as described in Table 1.

Since different stainings highlight different structures, the pathologist only assesses the lesion specified for the staining. Similarly, when performing annotation, glomeruli are labeled only with the staining-specified lesion shown in Table 1. NOA means none of all staining-specified lesions. To achieve accurate validation during training and evaluation after training, the validation and test datasets were annotated by two advanced pathologists to guarantee correct labeling. For much larger training datasets, the interns perform the annotation.

### B. PROBLEM SETTINGS

In the classification of glomeruli with noisy labels, our goal is to find a mapping function $f$ between the pathological images space $\mathcal{X}$ and the annotated set with noise $\hat{\mathcal{Y}}$. The ground truth space $\mathcal{Y}$ without noisy labels is potentially available but practically inaccessible to us due to insignificance in the lesion, insufficient resolutions in images, annotation workload and subjectivity. Traditionally, to optimize model $f$ with clean labelis obtained, a supervised loss function $\mathcal{L}$ is designed to penalize the difference between the model prediction $f(x)$ and the ground truth $y$, as the optimal model is

$$f^* = argmin\mathbb{E}_{(x,y)\ (\mathcal{X},\mathcal{Y})}[\mathcal{L}(f(x), y)]. \tag{1}$$

The difference between $\hat{\mathcal{Y}}$ and $\mathcal{Y}$ bring this method the semantic ambiguity, thus falling into overfitting and mislearning. Therefore, we need to identify a new optimization way to find a mapping function $f$, so that training with $\{\mathcal{X}, \hat{\mathcal{Y}}\}$ yields an appreciable generalization performance on a clean test set.

### C. THE OVERVIEW OF THE PROPOSED MODEL

The overview of the model is shown in Fig. 2. Before training, Mask R-CNN is used to detect the glomeruli on the WSIs followed by the annotations. The labels in the training dataset are considered to consist of clean and noisy labels, which are updated dynamically during training. The annotation noise interferes with the network and shows excessive uncertainty in the prediction of the image. Therefore, noisy label discriminator (NLD) performs the recognition of noises by calculating sample uncertainties, where samples with higher uncertainties are considered carrying noisy labels.

During training, augmentation $A$ transforms one image to generate two different images, and these two images are fed into Siamese networks that share weights. One of the networks computes a supervised classification loss on clean data, which drives the feature extractor to biasly encode the pathological features associated with each lesion.

In addition to the supervised classification loss, the Siamese network computes an unsupervised contrast loss. It uses a mapping layer to project two sets of homologous images into the same feature space.Both images originating from the same image are consistent in terms of the semantic information of the glomerular lesion. This implies that the distances between their features in the new feature space should be relatively close. This regularization encodes the same semantic features closer to each other in the projection space and vice versa, motivating the feature extractor to learn salient features. It is computed using an unsupervised contrastive loss for both clean and noisy data.

In addition to contrastive learning, a teacher-student model is used to perform consistency regularization. The teacher model is obtained by exponential moving averaging (EMA) from one of the Siamese models called student model. It prevents the student from getting stuck in local optima and provides more stable predictions. This is achieved by constraining the agreement between student and teacher predictions.

### D. MODEL DETAILS

#### 1) NOISY LABEL DISCRIMINATOR

The space $\hat{\mathcal{Y}}$ of the dataset constructed from clinical data carries a lot of label noises consisting of a potentially clean annotation space $\hat{\mathcal{Y}}_c$ and a noisy annotation space $\hat{\mathcal{Y}}_n$ due to insignificant lesion structure variation, insufficient resolution of the microscopic images, heavy and subjective annotation. Under the traditional paradigm, the training can be standardized by minimizing the difference between the prediction $f(x)$ and $y \in \hat{\mathcal{Y}}_c$, so that it captures lesion-related features. However, it is harmful to use $y \in \hat{\mathcal{Y}}_d$ directly for supervised

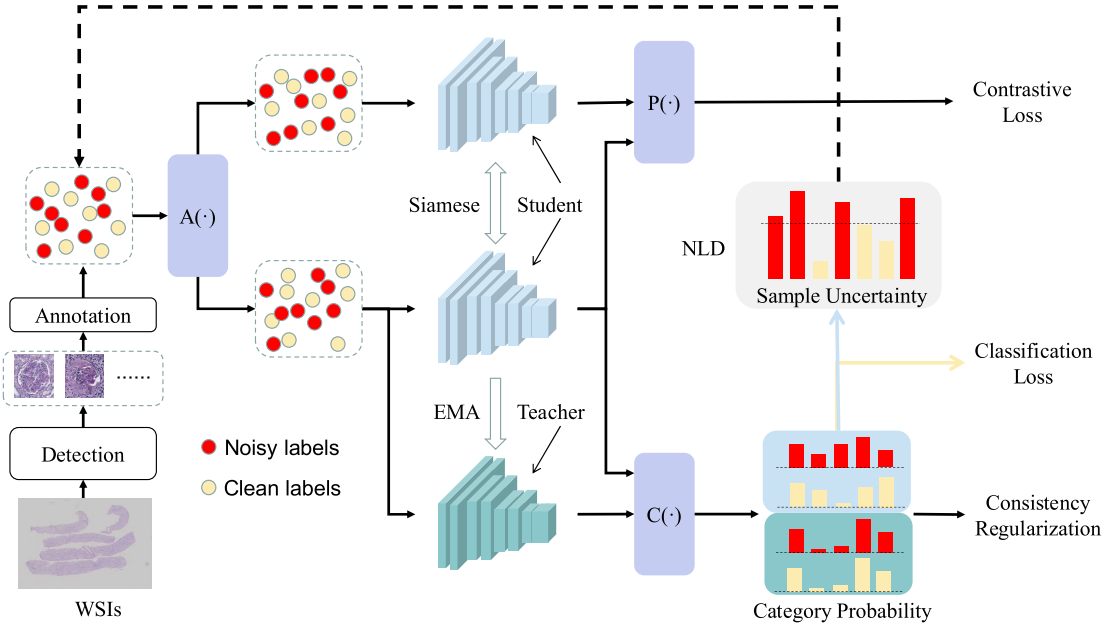



**FIGURE 2.** The overview of the proposed model. The detection and annotation are performed before the model training.

training, which can severely mislead the optimization of the network and make the network fall into overfitting or local optimum.

To address the above issues, NLD is designed to further distinguish noisy samples from clean ones. In general, lesions cause structural changes in the glomeruli. A typical lesion structure will help to give a higher confidence in the annotation, while an atypical one tends to increase subjectivity. For deep neural networks, the uncertainty value of the typical lesion structure prediction tends to be lower, and vice versa. Therefore, we use uncertainty as a metric to distinguish clean samples from noisy ones. The operation of NLD consists of three steps. First, the uncertainty is calculated for each sample. The sample uncertainty was defined as the maximum information entropy of the prediction of each category:

$$U_{x_i} = \max_{1 \le c \le C} -(f_c(x_i) log f_c(x_i) + (1 - f_c(x_i)) log(1 - f_c(x_i))), \tag{2}$$

where $i$ and $j$ are the index of training samples and categories respectively, and $C$ mean the total categories of this stain. Second, samples with the largest uncertainty of $p\%$ are selected as noisy samples. Third, noisy samples with stain-related lesion labels are returned as clean samples, since the number of positive labels is much smaller than the number of negative labels. After each epoch of training, the obtained grouping information of clean and noisy samples is used in the next round, so the clean and noisy samples are updated dynamically.

### 2) CONTRASTIVE LEARNING

After obtaining clean and noisy samples, we employ unsupervised contrast loss to exploit the rich image features of noisy samples to improve the feature extraction ability of the network.The computation process of contrast loss is shown in Figure 3. In one batch, image $x_i$ is transformed into $x_{i1} \in \mathbb{R}$ and$x_{i2}$ by random augmentation $A$. $x_{i1}$ and $x_{i2}$ are fed into the encoder of the Siamese network, and finally go through the projection layer $P$ to obtain embeddings $z_{i1} \in \mathbb{R}^{128}$ and $z_{i2} \in \mathbb{R}^{128}$. Another image $x_j$ in the same batch goes through exactly the same steps to finally obtain $z_{j1} \in \mathbb{R}^{128}$ and $z_{j2} \in \mathbb{R}^{128}$. $P$ is a multilayer perceptron (MLP), which contains a hidden layer, ReLU activation layer and fully connected (FC) layer. $P$ encodes the 2048-dimensional features obtained from the encoder to 128 dimensions, and the vector contains the semantic information of the current sample in the underlying feature space.

We use the normalized temperature-scaled cross entropy loss [25] as the contrast loss. For each anchor $z_{i1}$ in a batch, its homologue $z_{i2}$ is the positive sample carrying the same semantic information, while all other samples $z_k$ $(k \neq i)$ are negative samples. The contrast loss serves to maximize the similarity between the anchor and positive samples, while minimizing that between the anchor and negative samples. The formula of the loss function is

$$\mathbb{L}_{cl} = -log \frac{e^{sim(z_m, z_n)/\tau}}{\sum_{k=1}^{2N} \mathbb{1}_{k \neq m} e^{sim(z_m, z_n)/\tau}}, \tag{3}$$

where $sim$ means cosine similarity function and $N$ is the size of the batch.

### 3) CONSISTENCY REGULATIZATION

The student model performs supervised classification loss by NLD using clean labels thus achieving efficient feature extraction, while utilizes unsupervised contrastive loss to further enhance the ability to extract features. During training, clean and noisy data are dynamically updated, therefore the prediction of the student model fluctuates continuously.

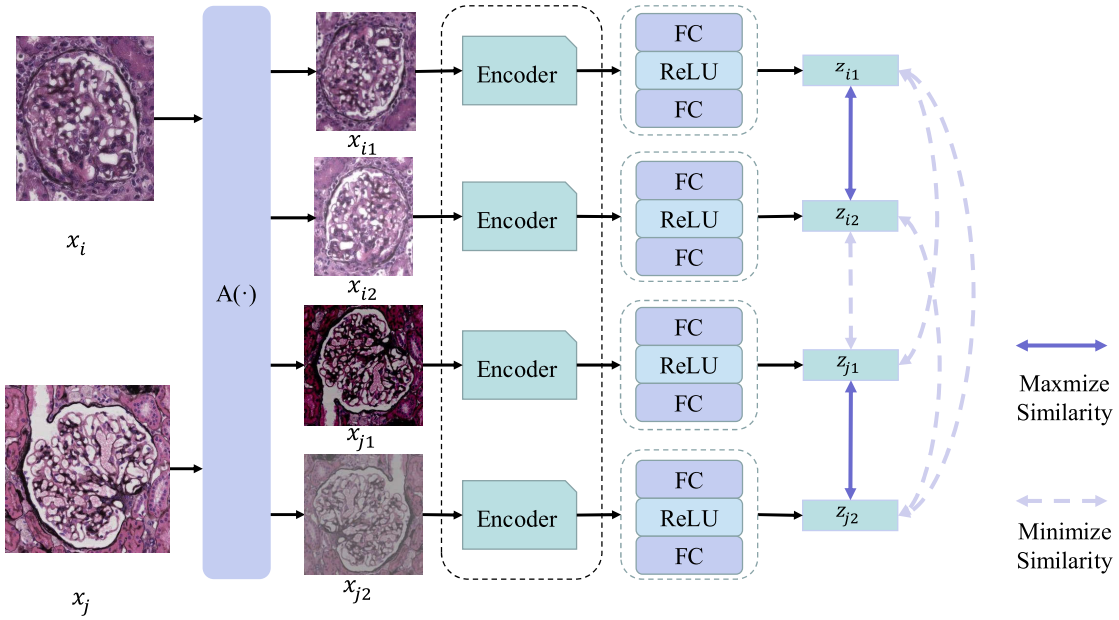



**FIGURE 3.** The overview of the contrastive learning. The encoders share the same weights.

To prevent instability and stucking in local optima in training, we take EMA of student as the teacher with more stable parameter updates to constrain the training.

During training, the teacher is a moving exponential weighted average of the student, updated by the formula:

$$\theta_t^{'} = \alpha\theta_{t-1}^{'} + (1-\alpha)\theta_t, \tag{4}$$

where $\theta^{'}$ and $\theta$ means the teacher model and student model respectively, $t$ means the iteration index. $\alpha$ is the smoothing parameter, controlling the smoothing of the parameters of the teacher. The teacher's prediction contains the student's history information and is therefore more stable. We define the consistency loss function to perform the regularization as

$$\mathbb{L}_{cr} = \left\| f_\theta(x) - f_{\theta^{'}}(x) \right\|_2. \tag{5}$$

During inference, the output of the student model are set as final predictions.

#### 4) BACKBONE AND OVERALL OPTIMIZATION OBJECTIVE

In this paper, pretrained ResNet-101 [31] was used as the backbone. The number of output nodes is set as the number of lesion categories of each staining. The binary cross-entropy loss function was used as

$$\mathbb{L}_c = \sum_i \sum_c^C -[y_{ic} log(p_{ic}) + (1-y_{ic}) log(1-p_{ic})], \tag{6}$$

where $p_{ic}$ means the predicted probability of $c$th category of the $i$ sample.

The overall optimization objective of the proposed method is

$$\min_\theta \mathbb{L}_c + w_{cl}\mathbb{L}_{cl} + w_{cr}\mathbb{L}_{cr}, \tag{7}$$

where $w_{cl}$ and $w_{cr}$ are the loss weights and $\theta$ is the model parameters.

## IV. RESULTS

### A. EXPERIMENTS SETUP

To validate the efficiency of the proposed method on three noisy glomerular datasets with different stainings (PAS, MT and PASM), we compare it with other methods for noisy label learning. Furthermore, we perform intensive experiments to demonstrate the usefulness of each module and discuss the effect of different sets of hyperparameters and implementations. At each trial, the model that performs best on the validation set is saved for inference.

#### 1) TRAINING SETTINGS

We train the proposed noisy label learning framework with Adam optimizer with a learning rate of 1e-4, using NVIDIA A100 GPUs and the batch size was set to 32. All images are resize to (224 pixles, 224 pixels), and are utilized random flip, rotation, scaling, random brightness and contrast shift and normalization as augmentation methods. NLD was run from beginning with the value of $p$ set as 10, which means the top $p\%$ uncertain samples are considered to be labeled incorrectly. The starting epoch of CL was set to 10, with the temperature $\tau$ set as 0.5 and $w_{cr}$ set as 0.1. The teacher model was introduced from the 15th epoch, where the maximun value of $\alpha$ was set to 0.9 and $w_{cr}$ is set as 0.1. $\alpha$ was defiend as

$$\alpha = \min(1 - \frac{1}{e_c - e_s + 1}, 0.9), \tag{8}$$

where $e_c$ and $e_s$ are the current and starting epoch, respectively.

#### 2) EVALUATION METRICS

During training, the model with the highest average $f_1$-score of each category on the validation set is considered the best

**TABLE 2. The comparison of different methods.**

| Methods | $mean - f_1$ | | | $mean - acc$ | | |
|---|---|---|---|---|---|---|
| | PAS | MT | PASM | PAS | MT | PASM |
| Baseline | 78.04 | 71.51 | 69.12 | 93.37 | 93.42 | 94.02 |
| RobustLoss | 75.16 | 77.57 | 70.03 | 93.37 | **96.82** | **96.82** |
| NoiseDrop | 78.50 | 77.39 | 70.31 | **94.60** | 95.23 | 94.51 |
| LabelCorrection | 79.79 | **77.70** | 70.92 | 94.19 | 95.92 | 94.11 |
| Proposed | **80.71** | 77.30 | **71.14** | 94.34 | 95.74 | 95.42 |

model. In the inference process, the average $f_1 - score$ and accuracy of each lesion on the test set are used as the evaluation metrics. We define mean $f_1 - score$ as $mean - f_1$ which is calculated as

$$mean - f_1 = \frac{1}{C}\sum_{c=0}^{C}\frac{2P_cR_c}{P_c + R_c}, \tag{9}$$

where $P_c$ and $R_c$ are the precision and recall of $c$th category, respectively. The $mean - acc$ are the average accuracy of all categories.

### B. COMPARISON TO OTHER METHODS

To verify the superiority of our method, we compare the following methods:

- Baseline: Pretrained ResNet-101.
- RobustLoss: Baseline, equipped by symmetric cross-entropy (SCE) loss which is robust to annotation noise [4].
- NoiseDrop: Two pretrained ResNet-101s with weight sharing performing co-teaching, with noisy samples discarded [7].
- LabelCorrection: Pretrained ResNet-101 with correction for noisy labels [32].
- Proposed: The proposed method.

The results are presented in Table 2, with the best performance indicated in bold letters. The proposed method shows the best $mean - f_1$ on both PAS and PASM, and was comparable to LabelCorrection on MT. Due to data imbalance, $mean - acc$s is about the same and our method is better than Baseline significantly. As $mean - f_1$ is fair for unbalanced datasets, our model is superior to other methods.

To further demonstrate the enhancement, we list $f_1 - scores$ of each lesion using Baseline and proposed method in Fig. 4. In summary, the proposed method achieves accurate results, where 25% (4/16) of the lesions had $f_1 - score$ above 85%, 43.75% (7/16) had $f_1 - score$ above 80%, and 75% (12/16) had $f_1 - score$ at or above 70%. Among them, GS is with the highest $f_1 - score$ for three stainings (all around 90%). $f_1 - score$ of MH, PAS(-) and Cre in PAS is over 80%, but $f_1 - score$ of SS is lower. $f_1 - score$ of SFN and Cre in MT is around 70%. $f_1 - score$ of Cre in PASM reaches over 85%, while IS, SS, and SFN are lower. From a comparative point of view, the proposed method achieves improvements over Baseline. The huge improvements include PAS-PAS(-) (+5.16%), PAS-Cre (+4.97%), MT-SFN (+5.66%), MT-Cre (+6.31%), PASM-EP (+5.00%) and PASM-Cre (+5.50%). Additionly, there are significant gains for PAS-SS (+2.72%), MT-GS (+2.43%), PASM-MP (+2.43%), and GS (+2.83%).

### C. VISUALIZATION

To demonstrate the evidence the model locate to recognize lesions, we use Grad-CAM [33] to visualize the feature maps, depending on which model makes the recognition. Fig. 5 illustrates that by discriminating and processing noises, the proposed model focuses on the fine-grained lesion-related regions. A detailed description is given below.

- For PAS, the model identified mesangial regions with dense mesangial cells to recognize MH in 5.a. The PAS(-) in 5.b is predicted with the most PAS-negative mesangial matrix found. In 5.c and 5.d, SS and Cre are recognized; moreover, the model precisely focuses on the segment sclerosis and proliferation of glomerular parietal epithelial cells.
- For MT, SFN in 5.e is recognized by identifying the fibrinoid necrosis of capillary wall. For Cre in 5.f and GS in 5.g, the model pays attention to proliferating parental epithelial cells and global sclerosis. In 5.h, the model did not capture valid features for NOA.
- For PASM, the model focuses on the whole of the crumpled glomerulus for IS in 5.i. In 5.k, the MN is predicted by identifying vacuolar degeneration of basement membrane. In 5.k and 5.l, the model identifies most areas where vacuolar degeneration of basement membrane and where the endocapillary proliferation exists, for MN and EP, respectively.

### D. ABLATION STUDY

In this section, we perform extensive ablation experiments to demonstrate the effectiveness of the proposed module, and in addition, we discuss the role of hyperparameters and ensemble inference methods.

#### 1) ABILITIES OF MODULES

To demonstrate the significance of each module, we compare the models presented in Table 3. ✓means the use of modules. As in rows 2-4, NLD, CL, and CR all result in improvements compared to Baseline in row 1. This illustrates the effectiveness of these modules in enhancing discriminating noise discrimination and feature extraction. However, mixing modules makes the results complicated. For example, Baseline+NLD+CL performs better than Baseline+NLD on most metrics, while Baseline+NLD+CR performs worse than Baesline+CR. As proposed, the combination of NLD, CL,

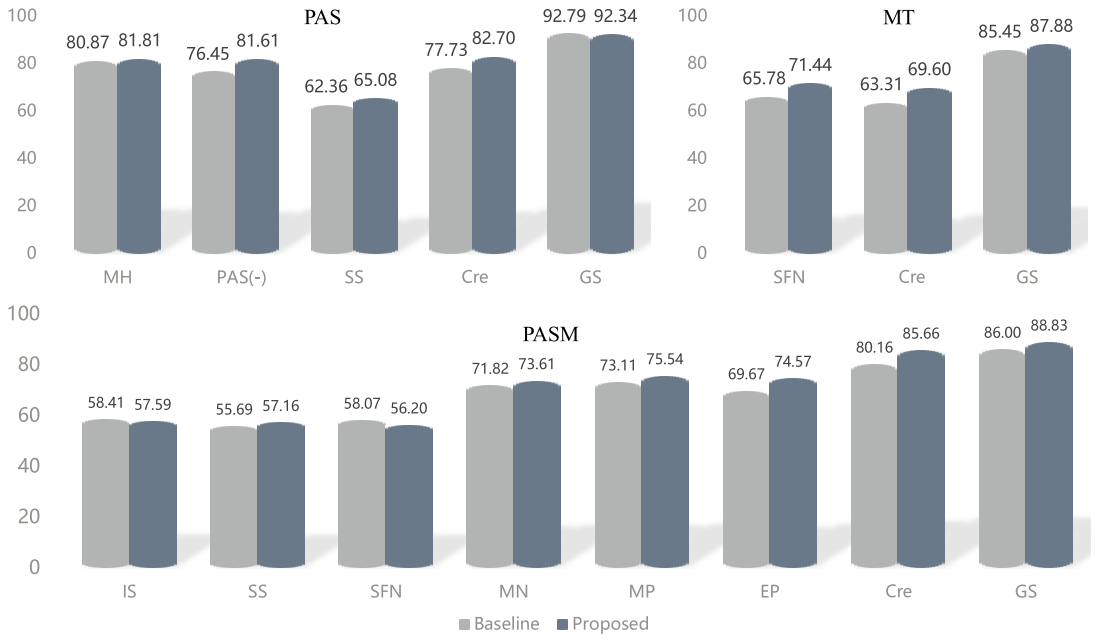



**FIGURE 4.** The results of the proposed method on all categories of three stainings.

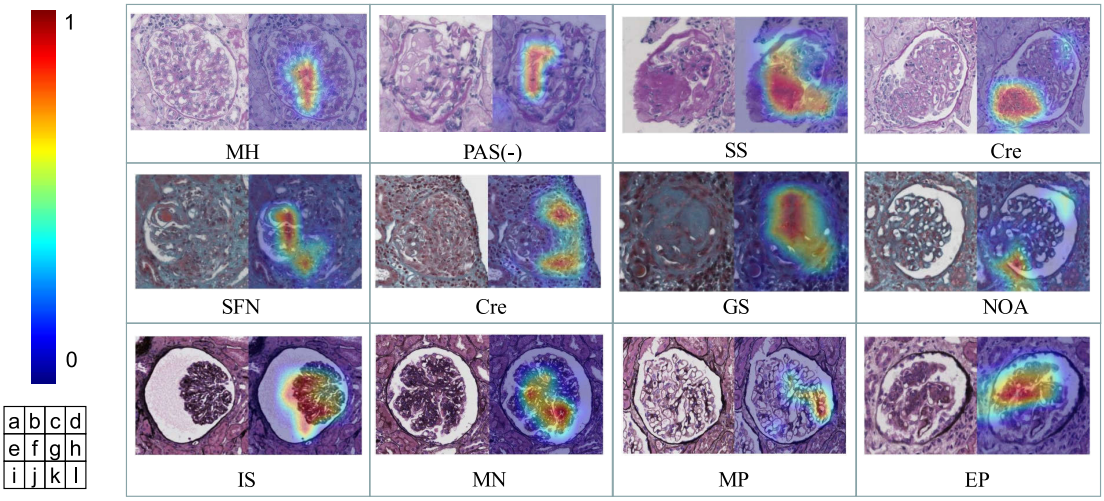



**FIGURE 5.** Visualization of Grad-CAM of the proposed model. Row 1-3 are PAS, MT and PASM, respectively. The color bar on the left represents the attention value.

and CR achieves the best results, beating other methods on all metrics.

### 2) IMPLEMENTATION OF NLD

To verify the effectiveness of the NLD, we compare the following models: (1) Baseline; (2) Baseline + Loss-based NLD (L-NLD) [34]; (3) Baseline + Uncertainty-based NLD with $p\% = 10\%$ (U-NLD-$p = 10\%$); (4) Baseline + Uncertainty-based NLD with $p\% = 30\%$ (U-NLD-$p = 30\%$); (5) Baseline + Uncertainty-based NLD with $p\% = 50\%$ (U-NLD-$p = 50\%$). L-NLD determines the noisy labels based on the classification loss, which is fit and calculated by Gaussian mixture model (GMM), with samples with the largest loss of 10% considered as noisy ones. The results shown in Table 4 illustrate U-NLD-p=10% is the best. Compared to Baseline, L-NLD gain improved only on PAS, while U-NLD outperforms Baseline with p=10% or 30%. For a too large value of p as 50%, the performance of U-NLD is worse.

### 3) INFLUENCE OF TEMPERATURE

To investigate the effects of temperature $\tau$ in the contrast loss, we compare the performance under different values of 0.01, 0.1 and 1. The results are illustrated in Table 5. The model with *tau* set to 0.5 performs the best, with both the $mean{-}f_1$ on PAS and PASM better than the models with other threshold. For MT, model with *tau* set to 1 has the best highest $mean{-}f_1$. Generally, model with *tau* set to 0.1 performs worst for all three stainings.

**TABLE 3.** Comparison of different implementations.

| Baseline | NLD | CL | CR | $mean - f_1$ | | | mean-acc | | |
|---|---|---|---|---|---|---|---|---|---|
| | | | | PAS | MT | PASM | PAS | MT | PASM |
| √ | | | | 78.04 | 71.51 | 69.12 | 93.37 | 93.37 | 94.02 |
| √ | √ | | | 79.84 | 72.87 | 69.46 | 93.57 | 94.11 | 93.39 |
| √ | | √ | | 78.95 | 75.83 | 69.83 | 93.94 | 95.29 | 93.86 |
| √ | | | √ | 79.18 | 75.57 | 69.72 | 93.58 | 95.07 | 94.06 |
| √ | √ | √ | | 78.93 | 75.44 | 71.10 | 93.94 | 95.19 | 94.73 |
| √ | √ | | √ | 78.86 | 72.92 | 68.43 | 93.28 | 94.69 | 94.35 |
| √ | √ | √ | √ | **80.71** | **77.30** | **71.14** | **94.34** | **95.74** | **95.42** |

**TABLE 4.** Comparison of different implementation of NLD.

| Methods | $mean - f_1$ | | |
|---|---|---|---|
| | PAS | MT | PASM |
| Baseline | 78.04 | 71.51 | 69.12 |
| L-NLD | 79.69 | 69.86 | 68.55 |
| U-NLD-p=10% | **79.84** | **72.82** | 69.46 |
| U-NLD-p=30% | 79.49 | 72.65 | **69.51** |
| U-NLD-p=50% | 78.67 | 67.02 | 68.33 |

**TABLE 5.** Comparison of different temperatures.

| $\tau$ | $mean - f_1$ | | |
|---|---|---|---|
| | PAS | MT | PASM |
| 0.1 | 79.35 | 73.58 | 69.06 |
| 0.5 | **80.55** | 74.72 | **71.08** |
| 1 | 79.91 | **75.38** | 69.67 |

## V. DISCUSSION

Comprehensive information on glomerular lesions patterns is vital for kidney disease diagnosis. While deep learning can assist to achieve more objective, efficient and accurate recognition, its performance is sensitive to annotation noise. In this paper, we propose a novel noisy label learning model on glomerular images with annotation noise by discriminating noisy samples through NLD, and leveraging CL and CR to enhance the encoding of pathological features.

### A. SUPERIORITY OF THE PROPOSED METHOD

Experiments demonstrate the superiority of the proposed method. Baseline focuses too much on mislabeled labels, and RobustLoss limits the over-dependence of clean labels without directly handling the noises. For NoiseDrop, the discarded may have correct labels and carry discriminative information, so this method may lead to inadequate learning of features. As LabelCorrection may perform inaccurate label correction due to its dependence on models trained on noisy datasets, new noises will be introduced to hinder the optimization. For the proposed method, the noisy labels are discriminated using the uncertainty-based method. Unlike the above works, it treats noisy labels as unlabeled ones and employs CL and CR to enhance the encoding capability in an unsupervised manner. This not only excludes noisy labels for the supervised loss, but also facilitates the construction of decision boundaries.

Compared to Baseline, the proposed model brings gains on different lesions. The greatest gain is for PAS-PAS(-), PAS-Cre, MT-SFN, MT-Cre, PASM-EP and PASM-Cre, with all the improvements being more than 5%. For most of these lesions, a significant number of samples helps to perceive pathological features. However, the sensitivity to annotation noise causes the Baseline to be incorrectly optimized. However, for the proposed method, noisy labels can be separated, thus using large amounts of samples to improve under unsupervision.

For different lesions, our method shows different performance:

- Collectively, the best performance is for GS and Cre, with the $f_1 - scores$ are near 85% and even over 90%. The glomerulus of GS has almost no visible glomerular capillary loop cavity, with an uniform sclerotic state. In Cre, the parietal epithelial cells proliferated significantly in Bowman's cystic, accompanied by inflammatory cell infiltration and changes in capillary loops. With annotation noise discriminated by the proposed uncertainty-based NLD, the model uses clean labels to learn GS and Cre and enhances the ability to encode features by CL and CR. This results in optimal recognition and visualization shown in Fig. 4 and Fig. 5.
- Most lesions of PAS and PASM are well recognized. For PAS, $f_1 - score$ for MH and PAS(-) both reach over 80%. For MH, the width of mesangial area and number of mesangial cells are different in degrees (mild, moderate and severe). For these variable global morphologies of the glomeruli, the proposed model excludes annotation noises to extract degree-agnostic semantic features efficiently. For PAS(-), the mesangial matrix is increased and PAS staining is light (indicating the deposition of amyloid in the mesangial area). Our proposed method benefits from high signal-to-noise learning to capture such features to construct decision boundary, improving $f_1 - score$ by more than 5% compared to Baseline. For PASM, MN, MP and EP are well predicted by identifying finer structures of the basement membrane and capillaries, the basement membrane and cells in the capillaries shown in Fig.4. In the proposed model, the low-level semantic features represent finer pathological information that can be extracted under the supervision of the correct labels when noises are excluded. Therefore, the performance is considerably improved.

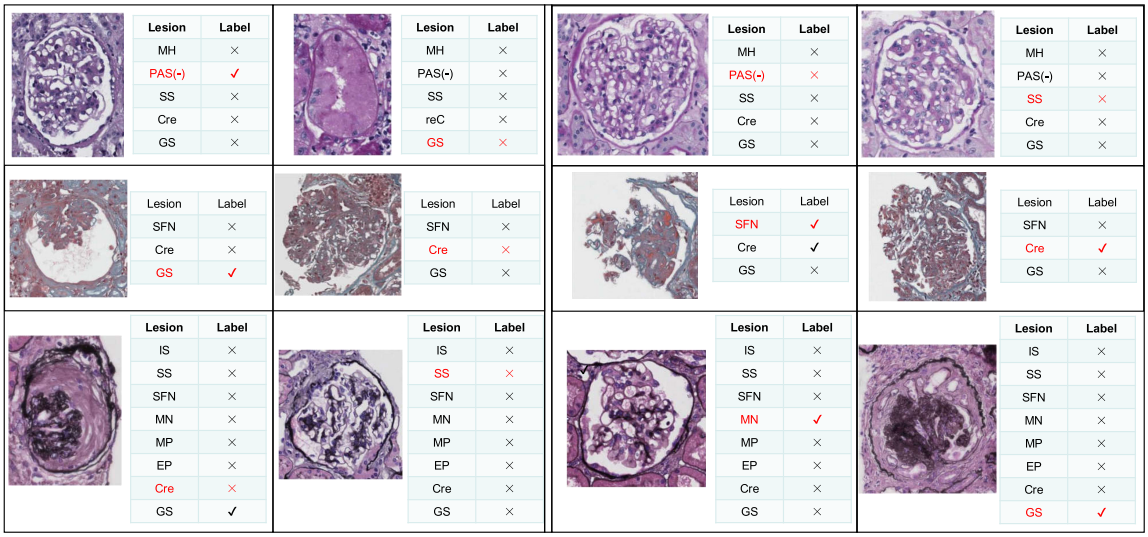



**FIGURE 6.** Examples of samples predicted as noisy ones by the proposed NLD. Colomn 1 and 2 refer to L-NLD, while column 3 and 4 refer to U-NLD. Colomn 1 and 3 mean the right NLD discrimination, while Colomn 2 and 4 mean the wrong discrimination. Red labels are predicted to be noisy.

- The recognition of PAS-SS, PASM-IS, PASM-SS and PASM-SFN is complicated by their progressive and high-resolution properties. Moreover, the sample number is too small to perceive lesion-related features. For SS, the difference in the proportion of affected capillary loops leads to a large variation in the degree of sclerotic homogeneity. A small fraction causes SS to be overwhelmed in other lesions, while a large fraction may be similar to GS. Similarly, for IS, the different degree of wrinkling leads to complexity. For SFN, it is necessary to identify that the basement membrane of capillary loop is broken under PASM staining. This requires not only a combination with electron microscopy or immunofluorescence, but also a considerable number of lesion samples. For one case, the glomeruli with SFN may only account for a small fraction of all glomeruli, so the number of SFN we obtained is limited. More images are needed to improve the perception of these lesions.

### B. INFLUENCE OF DIFFERENT IMPLEMENTATIONS OF NLD

In pathological datasets, NLD improves the generalization ability of the model. First, U-NLD performs better than L-NLD. For L-NLD, the correlation of classification loss with the probability of being noisy is weak, and the noisy labels degrades the GMM, while U-NLD is more label-independent. Second, larger $p\%$ (p=50) means half of the samples are dropped to lead to a loss of salient features, thus making recognition worse. Third, we visualize some noisy labels discriminated in Fig. 6. Both NLD can recognize noisy labels, while correct labels may be incorrectly discriminated. For handling these noises, NoiseDrop and LabelCorrection discard valuable lesion-related features or inevitably introduce new noises. However, the proposed model process in an unsupervised manner. It takes full advantage of the pathological features provided by the large number of images, avoids introducing new noise, and ultimately performs the best.

### C. INFLUENCE OF DIFFERENT TEMPERATURES OF CL

The role of temperature is to regulate the degree of attention the network pays to hard samples. Based on the experimental results, a smaller temperature is more focused on separating samples from similar hard samples, but may destroy valuable underlying semantic structures. When t is 0.1, the network pays much attention to hard samples, resulting in the destruction of the original semantic structure, thus performs the worst. When t is 1, the network does not gain from computation of unsupervised contrast loss. When t is 0.5, the network not only makes use of noisy samples to improve the performance of pathological feature extraction, but also avoids falling into local optimum, so it achieves the best.

## VI. CONCLUSION

Deep learning for glomerular lesion recognition is limited by annotation noise arising from different progression of pathological development of the glomeruli, insignificant lesion structures, and insufficient information under light microscopy. We propose a novel noisy label learning model that excludes noisy labels based on sample uncertainty and mining large-scale datasets to enhance the encoder to extract pathological features using CL and CR. On the PAS, MT, and PASM datasets, the proposed method outperformed in recognizing multiple lesions, and the effect of each module was demonstrated. Our work assists pathologists to perform efficient, objective and accurate recognition of glomerular lesions and lay foundation for the research of computer aided diagnosis for renal pathology.

There are still some shortcomings that need to be improved. Our work provides new ideas for developing algorithms on large-scale noisy datasets, but the performance on

complex glomerular lesion recognition needs a further boost. Larger datasets are required to observe a more diverse pathological features of lesions. In addition, the self-supervised pre-training with glomerular correlation will be performed to extract more salient pathological features, which will improve generalization.

## REFERENCES


[1] J. L. Kwek and T. Y. Kee, "World kidney day 2020: Advances in preventive nephrology," *Ann. Acad. Med.*, vol. 49, no. 4, pp. 175–179, 2020.

[2] A. Ghosh, H. Kumar, and P. S. Sastry, "Robust loss functions under label noise for deep neural networks," in *Proc. AAAI Conf. Artif. Intell.*, 2017, vol. 31, no. 1, pp. 1–15.

[3] Z. Zhang and M. Sabuncu, "Generalized cross entropy loss for training deep neural networks with noisy labels," in *Proc. Adv. Neural Inf. Process. Syst.*, vol. 31, 2018.

[4] Y. Wang, X. Ma, Z. Chen, Y. Luo, J. Yi, and J. Bailey, "Symmetric cross entropy for robust learning with noisy labels," in *Proc. IEEE/CVF Int. Conf. Comput. Vis. (ICCV)*, Oct. 2019, pp. 322–330.

[5] G. Pleiss, T. Zhang, E. Elenberg, and K. Q. Weinberger, "Identifying mislabeled data using the area under the margin ranking," in *Proc. Adv. Neural Inf. Process. Syst.*, vol. 33, 2020, pp. 17044–17056.

[6] S. Guo, W. Huang, H. Zhang, C. Zhuang, D. Dong, M. R. Scott, and D. Huang, "CurriculumNet: Weakly supervised learning from large-scale web images," in *Proc. Eur. Conf. Comput. Vis. (ECCV)*, 2018, pp. 135–150.

[7] B. Han, Q. Yao, X. Yu, G. Niu, M. Xu, W. Hu, I. Tsang, and M. Sugiyama, "Co-teaching: Robust training of deep neural networks with extremely noisy labels," in *Proc. Adv. Neural Inf. Process. Syst.*, vol. 31, 2018.

[8] J. Han, P. Luo, and X. Wang, "Deep self-learning from noisy labels," in *Proc. IEEE/CVF Int. Conf. Comput. Vis. (ICCV)*, Oct. 2019, pp. 5138–5147.

[9] K. Yi, G.-H. Wang, and J. Wu, "PENCIL: Deep learning with noisy labels," 2022, *arXiv:2202.08436*.

[10] C.-A. Weis, J. N. Bindzus, J. Voigt, M. Runz, S. Hertjens, M. M. Gaida, Z. V. Popovic, and S. Porubsky, "Assessment of glomerular morphological patterns by deep learning algorithms," *J. Nephrol.*, vol. 35, no. 2, pp. 417–427, Mar. 2022.

[11] E. Uchino, K. Suzuki, N. Sato, R. Kojima, Y. Tamada, S. Hiragi, and H. Yokoi, "Classification of glomerular pathological findings using deep learning and nephrologist–AI collective intelligence approach," *Int. J. Med. Inf.*, vol. 141, Jan. 2020, Art. no. 104231.

[12] Y. Nan, F. Li, P. Tang, G. Zhang, C. Zeng, G. Xie, Z. Liu, and G. Yang, "Automatic fine-grained glomerular lesion recognition in kidney pathology," *Pattern Recognit.*, vol. 127, Jul. 2022, Art. no. 108648.

[13] C.-K. Yang, C.-Y. Lee, H.-S. Wang, S.-C. Huang, P.-I. Liang, J.-S. Chen, C.-F. Kuo, K.-H. Tu, C.-Y. Yeh, and T.-D. Chen, "Glomerular disease classification and lesion identification by machine learning," *Biomed. J.*, vol. 45, no. 4, pp. 675–685, Aug. 2022.

[14] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2016, pp. 770–778.

[15] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," 2014, *arXiv:1409.1556*.

[16] A. Dosovitskiy, L. Beyer, A. Kolesnikov, D. Weissenborn, X. Zhai, T. Unterthiner, M. Dehghani, M. Minderer, G. Heigold, S. Gelly, J. Uszkoreit, and N. Houlsby, "An image is worth 16×16 words: Transformers for image recognition at scale," 2020, *arXiv:2010.11929*.

[17] Z. Liu, Y. Lin, Y. Cao, H. Hu, Y. Wei, Z. Zhang, S. Lin, and B. Guo, "Swin transformer: Hierarchical vision transformer using shifted windows," in *Proc. IEEE/CVF Int. Conf. Comput. Vis. (ICCV)*, Oct. 2021, pp. 10012–10022.

[18] G. Campanella, M. G. Hanna, L. Geneslaw, A. Miraflor, V. W. K. Silva, K. J. Busam, E. Brogi, V. E. Reuter, D. S. Klimstra, and T. J. Fuchs, "Clinical-grade computational pathology using weakly supervised deep learning on whole slide images," *Nature Med.*, vol. 25, no. 8, pp. 1301–1309, 2019.

[19] M. Y. Lu, T. Y. Chen, D. F. Williamson, M. Zhao, M. Shady, J. Lipkova, and F. Mahmood, "AI-based pathology predicts origins for cancers of unknown primary," *Nature*, vol. 594, no. 7861, pp. 106–110, 2021.

[20] C. Northcutt, L. Jiang, and I. Chuang, "Confident learning: Estimating uncertainty in dataset labels," *J. Artif. Intell. Res.*, vol. 70, pp. 1373–1411, Apr. 2021.

[21] M. Toneva, A. Sordoni, R. Tachet des Combes, A. Trischler, Y. Bengio, and G. J. Gordon, "An empirical study of example forgetting during deep neural network learning," 2018, *arXiv:1812.05159*.

[22] J. Huang, L. Qu, R. Jia, and B. Zhao, "O2U-Net: A simple noisy label detection approach for deep neural networks," in *Proc. IEEE/CVF Int. Conf. Comput. Vis. (ICCV)*, Oct. 2019, pp. 3326–3334.

[23] C. Zhu, W. Chen, T. Peng, Y. Wang, and M. Jin, "Hard sample aware noise robust learning for histopathology image classification," *IEEE Trans. Med. Imag.*, vol. 41, no. 4, pp. 881–894, Apr. 2022.

[24] K. He, H. Fan, Y. Wu, S. Xie, and R. Girshick, "Momentum contrast for unsupervised visual representation learning," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2020, pp. 9729–9738.

[25] T. Chen, S. Kornblith, M. Norouzi, and G. Hinton, "A simple framework for contrastive learning of visual representations," 2020, *arXiv:2002.05709*.

[26] O. Henaff, "Data-efficient image recognition with contrastive predictive coding," in *Proc. Int. Conf. Mach. Learn.*, 2020, pp. 4182–4192.

[27] A. Tarvainen and H. Valpola, "Mean teachers are better role models: Weight-averaged consistency targets improve semi-supervised deep learning results," in *Proc. Adv. Neural Inf. Process. Syst.*, vol. 30, 2017.

[28] Q. Xie, M.-T. Luong, E. Hovy, and Q. V. Le, "Self-training with noisy student improves ImageNet classification," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2020, pp. 10687–10698.

[29] Y. Han, S. K. Roy, L. Petersson, and M. Harandi, "Learning from noisy labels via discrepant collaborative training," in *Proc. IEEE Winter Conf. Appl. Comput. Vis. (WACV)*, Mar. 2020, pp. 3169–3178.

[30] K. He, G. Gkioxari, P. Dollár, and R. Girshick, "Mask R-CNN," in *Proc. IEEE Int. Conf. Comput. Vis.*, Oct. 2017, pp. 2961–2969.

[31] J. Deng, W. Dong, R. Socher, L.-J. Li, K. Li, and L. Fei-Fei, "ImageNet: A large-scale hierarchical image database," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2009, pp. 248–255.

[32] J. Lee, H. Lim, and K.-S. Chung, "CLC: Noisy label correction via curriculum learning," in *Proc. IEEE Symp. Ser. Comput. Intell. (SSCI)*, Dec. 2021, pp. 1–7.

[33] R. R. Selvaraju, M. Cogswell, A. Das, R. Vedantam, D. Parikh, and D. Batra, "Grad-CAM: Visual explanations from deep networks via gradient-based localization," in *Proc. IEEE Int. Conf. Comput. Vis. (ICCV)*, Oct. 2017, pp. 618–626.

[34] J. Li, R. Socher, and S. C. Hoi, "DivideMix: Learning with noisy labels as semi-supervised learning," in *Proc. Int. Conf. Learn. Represent.*, 2020.

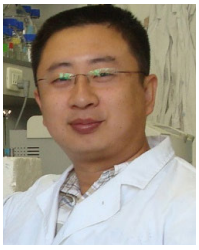

**JING LI** received the bachelor's degree in clinical medicine and the M.D. degree from Fourth Military Medical University, Xi'an, China, in 1998 and 2006, respectively. She finished the resident training with the Department of Pathology, Xijing Hospital, and got an associate professor position, in 2014. She was trained as a Dr. Fernand M. Lai's Fellow, in 2013, who is a worldwide famous Renal Pathologist of the Department of Anatomical and Cellular Pathology, Prince of Wales Hospital, The Chinese University of Hong Kong. She was also trained as Dr. Helmut G. Rennke's Fellow in 2015, who is a worldwide famous Renal Pathologist of the Department of Pathology, Brigham and Women's Hospital, Boston. She is mainly engaged in the research of vaccine for the hemorrhagic fever with renal syndrome. She has published more than ten articles, hosted and completed one National Natural Science Foundation of China Project.

**QIMING HE** received the B.S. degree in mechanical engineering from the University of Science and Technology Beijing, China, in 2020. He is currently pursuing the M.S. degree in biomedical engineering with Tsinghua University, Beijing. His research interests include AI-based medical image analysis and computational pathology. He has published academic papers in several international journals and conferences, and applied for national invention patents, involving histopathological diagnosis and immunohistochemical quantitative analysis of breast cancer, pathological diagnosis of liver biopsy, and virtual staining.

**YIQING LIU** received the B.S. degree in electrical engineering from North China Electric Power University, Beijing, China, in 2017, and the M.E. degree in biomedical engineering from Tsinghua University, Beijing, in 2020, where he is currently pursuing the Ph.D. degree in biomedical engineering. His research interests include deep learning and its applications in histopathological image analysis. He has published academic papers and made oral reports in many international journals and conferences, involving cell detection of histopathological sections and histopathological diagnosis of breast cancer.

**YANXIA WANG** received the Ph.D. degree in pathology and pathophysiology from Air Force Medical University, in 2018. Since 2019, she has been an Assistant Professor with the Pathology Department, Air Force Medical University. Since 2020, she has been a Master's Supervisor and engaged in renal pathological diagnosis and pulmonary circulation and pulmonary blood vessel research. She has published more than ten articles and completed one National Natural Science Foundation grant.

**TIAN GUAN** received the Ph.D. degree in biomedical engineering from Tsinghua University. He has conducted research in the field of biomedical engineering for many years, focusing on medical processes and devices, mainly including speech processing algorithms, auditory system modeling detection, and rehabilitation hearing aid systems. He has also attempted industrialization transformation in several projects, such as otoacoustic emission detector and cochlear implant devices. As the project leader, he has undertaken more than ten research projects, such as the National Natural Science Foundation of China, the Guangdong Ministry of Education Industry-University-Research Combined Project, and the Shenzhen Discipline Layout Project. He has published more than 90 high-level articles. He has also been successively awarded as local level/ reserve level high-level leading talents by Shenzhen.



**JING YE** received the bachelor's and M.D. degrees from Fourth Military Medical University, Xi'an, China, in 1997 and 2003, respectively. He was a Visiting Scholar and a Postdoctoral Fellow with the Laboratory of Academician Peng Li, The Hong Kong University of Science and Technology (HKUST), from June 2004 to September 2005, and Tsinghua University, from September 2005 to September 2007, exploring fat metabolism in the liver. He is currently a Professor, a Doctoral Supervisor, and the Deputy Director of the Department of Pathology, Basic Medical Science Academy, Air Force Medical University, and the Pathology Department, Xijing Hospital. He is mainly engaged in the research of diseases related to glucose and lipid metabolism disorders. Furthermore, he has completed postdoctoral research in the team of Prof. Tak W. Mak with the Ontario Cancer Institute (OCI)/Princess Margaret Hospital (PMH), Canada, focusing on the role of IDH mutations in tumor metabolic reprogramming. In recent years, he has hosted eight National Natural Science Foundation of China Youth projects and general projects. He has published more than 20 science citation index papers as the corresponding author or the first author, such as *Cell Metabolism*, in 2009, *Hepatology*, in 2015, *Proceedings of the National Academy of Sciences of the United States of America*, in 2017, and *Blood*, in 2021.

**YONGHONG HE** received the Ph.D. degree in biophysics from South China Normal University, China, in 2002. He was a Postdoctoral Researcher with Cranfield University, U.K., from 2002 to 2004. Currently, he is a Professor with the Graduate School at Shenzhen, Tsinghua University, China. He has been engaged in interdisciplinary scientific research of biomedical optical methods, technologies, applications and instruments for many years, presided over dozens of projects, such as the National Natural Science Foundation (ten projects), the National Science and Technology Plan, the provincial and municipal science and technology plan, and the enterprise science and technology development, and has successively carried out the digital intelligence of medical microscopic images, optical coherence tomography (OCT) technology, optical weak measurement, and unmarked biomolecular interaction sensing.

**ZHE WANG** received the Graduate and M.D. degrees from Fourth Military Medical University, Xi'an, China, in 1993 and 2000, respectively. He finished the resident training with the Department of Pathology, Xijing Hospital, he got a professor position, in 2011. He has been nominated as the Chairperson, since 2013. He was a Postdoctoral Researcher with the Hormel Institute, University of Minnesota, USA, from 2005 to 2007. He was trained as a Dr. John K. C. Chan's Fellow, in 2012, who is a worldwide famous Pathologist of the Department of Pathology, Elizabeth Hospital, Hong Kong. His research interests include the key kinase biology of esophageal carcinoma and artificial intelligent pathology. He is an expert pathologist in soft tissue, lymphoma and genitourinary pathology. He has published more than 180 peer reviewed articles. He is an Associated Editor of *Pathology-Research and Practice* journal.