Stochastic Label Refinery: Toward Better Target Label Distribution

Xi Fang, Jiancheng Yang, Bingbing Ni

Shanghai Jiao Tong University

25th INTERNATIONAL CONFERENCE ON PATTERN RECOGNITION
Milan, Italy 10 | 15 January 2021

Introduction

Background

- Most studies concentrate on data augmentation, model structure or loss function to improve the model ability in deep supervised learning, rather than target label distribution;
- Soft label may better than hard label: one-hot or other form of hard label leads to over-confidence; Label smoothing and Mixup show the excellent performance of soft label;
- Better labels can reduce the impact of noise label or long-tail data which are common in practical application;

Problem

How to get even better target label distributions?

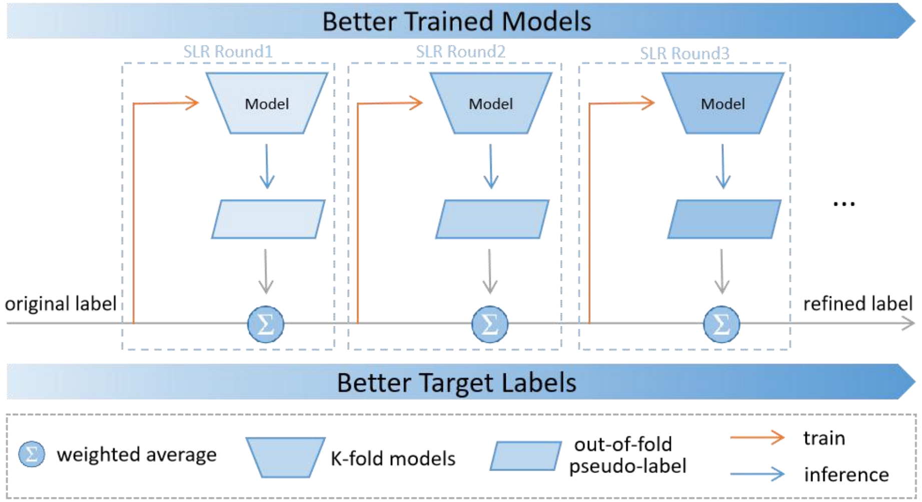
Contribution

- We point out the shortcomings of hard label distribution: (a) the risk of over-confidence, (b) easily affected by noise annotation and (c) lost intra-class and inter-class association;
- We propose a new regularized strategy of generating soft target label distribution: Stochastic Label Refinery (SLR);

Categories		Sky	Cloud	Sea	Plant	Others
Hard Label	one-hot	1	0	0	0	0
	multi-hot	1	1	0	1	0
Soft Label	label smoothing	0.9	0.025	0.025	0.025	0.025
	ideal distribution	0.45	0.1	0.05	0.35	0.05

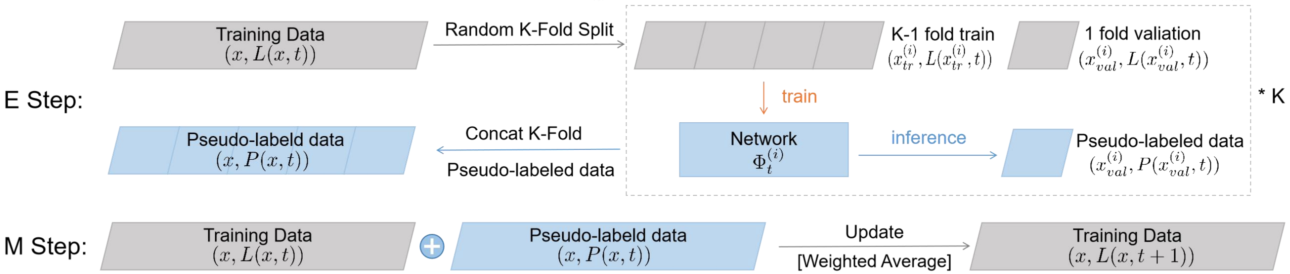
Methodology

Overall training pipeline of SLR



Details of each SLR round

1. randomly k-fold split the data to train k models;
2. inference them in k-fold validation data respectively to get out-of-fold(oof) pseudo-label; (E step)
3. use oof pseudo-label to refine the target label distribution by weighted average; (M step)
4. repeat from 1.

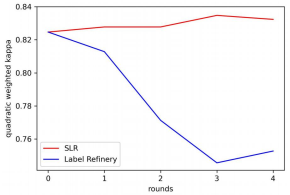


Results

SOTA in DeepDR Diabetic Retinopathy Dataset

Method	Quadratic Weighted Kappa
Baseline (w/o tricks)	0.8036±0.0214
Baseline (w/ tricks)	0.8247±0.0125
SWA [34]	0.8119±0.0234
OHEM [20]	0.8061±0.0174
Knowledge Distillation [27]	0.8128±0.0100
Label Refinery [16]	0.7527±0.0152
Stochastic Label Refinery	**0.8348±0.0053**

Method	Public Test	Private Test
Ours	**0.9303**	**0.9215**
Team1	0.9262	0.9211
Team2	0.9232	0.9097
Team3	0.9202	0.8946
Team4	0.9088	0.8890



Performance on Plant Pathology Dataset

Method	Top-1 Accuracy	Average AUC
Baseline	0.9676±0.0056	0.974±0.000
Focal Loss [8]	0.9665±0.0064	0.968±0.000
OHEM [20]	0.9670±0.0071	0.974±0.000
Label Smoothing [12]	0.9736±0.0106	0.974±0.000
Knowledge Distillation [27]	0.9731±0.0090	0.973±0.000
Label Refinery [16]	0.9720±0.0079	0.961±0.000
Stochastic Label Refinery	**0.9747±0.0083**	**0.976±0.000**