

Рецензия на **Sharpness-Aware Minimization for efficiently improving generalization (P. Foret, A. Kleiner, H. Mobahi, B. Neyshabur)** от Рахматуллина Рамазана(182)

Статья предлагает новый метод регуляризации модели для улучшения обобщающей способности. Идея заключается в том, что хороший минимум должен иметь гладкую поверхность вокруг себя, чтобы функция не изменялась слишком сильно при слабом изменении параметров.

В попытке теоретически обосновать предложенную идею, авторами метода была получена некоторая оценка сверху на функцию потерь по распределению, которую для простоты понимания было решено оценить как стандартную L2 регуляризацию плюс SAM-loss - максимальное отклонение функции потерь в окрестности точки. Далее путем несложных вычислений авторы статьи упростили SAM-loss до вычисления градиента в двух точках.

Интересным моментом является то, что полученная оценка имеет удобную геометрическую интерпретацию, причем вычисление оценки возможно довольно просто реализовать методами популярных фреймворков, замедляя время тренировки в два раза.

Главным результатом статьи авторы считают практическое обоснование применения их метода, поскольку они смогли получить новые state of the art результаты в известных задачах компьютерного зрения. В целом можно сказать, что в статье проведен достаточный практический анализ работы метода для заключения о его пригодности. В статье также упоминается применение к label noise, где SAM хорошо себя показал. Метод также был применен для задач finetuning, авторы заявляют о получении SOTA результатов. Однако в данном случае возникают вопросы о методах оценки SOTA, ведь судя по их таблице baseline уже превосходит SOTA в некоторых задачах.

Dataset	EffNet-b7 + SAM	EffNet-b7	Prev. SOTA (ImageNet only)	EffNet-L2 + SAM	EffNet-L2	Prev. SOTA
FGVC_Aircraft	6.80 \pm 0.06	8.15 \pm 0.08	5.3 (TBMSL-Net)	4.82 \pm 0.08	5.80 \pm 0.1	5.3 (TBMSL-Net)
Flowers	0.63 \pm 0.02	1.16 \pm 0.05	0.7 (BiT-M)	0.35 \pm 0.01	0.40 \pm 0.02	0.37 (EffNet)
Oxford_IIT_Pets	3.97 \pm 0.04	4.24 \pm 0.09	4.1 (Gpipe)	2.90 \pm 0.04	3.08 \pm 0.04	4.1 (Gpipe)
Stanford_Cars	5.18 \pm 0.02	5.94 \pm 0.06	5.0 (TBMSL-Net)	4.04 \pm 0.03	4.93 \pm 0.04	3.8 (DAT)
CIFAR-10	0.88 \pm 0.02	0.95 \pm 0.03	1 (Gpipe)	0.30 \pm 0.01	0.34 \pm 0.02	0.63 (BiT-L)
CIFAR-100	7.44 \pm 0.06	7.68 \pm 0.06	7.83 (BiT-M)	3.92 \pm 0.06	4.07 \pm 0.08	6.49 (BiT-L)
Birdsnap	13.64 \pm 0.15	14.30 \pm 0.18	15.7 (EffNet)	9.93 \pm 0.15	10.31 \pm 0.15	14.5 (DAT)
Food101	7.02 \pm 0.02	7.17 \pm 0.03	7.0 (Gpipe)	3.82 \pm 0.01	3.97 \pm 0.03	4.7 (DAT)
ImageNet	15.14 \pm 0.03	15.3	14.2 (KDforAA)	11.39 \pm 0.02	11.8	11.45 (ViT)

В конце статьи авторы вводят понятие m-sharpness, который они предлагают исследовать в последующих работах, ведь данное понятие по их заверениям является новой точкой зрения на регуляризацию путем изменения batch size. Однако эта часть статьи на мой взгляд является наименее подробной, поэтому к ней остаются вопросы.

В целом данная статья предлагает новый метод регуляризации, который хорошо себя показал на стандартных задачах компьютерного зрения. Его основным плюсом является применимость в произвольной дифференцируемой задаче, а также возможность применения вместе с другими методами регуляризации. У метода есть теоретическое обоснование, имеющее геометрическую интерпретацию, однако некоторые теоретические предположения(к примеру улучшение точности оценки SAM-loss при помощи гессиана) в статье не находят подтверждений. Статья также

имеет готовый код, адаптируемый к произвольным моделям в популярных фреймворках.

Я ставлю статье оценку 9, я уверен в оценке на 5/5.