Анализ смещения распределений в контрастивном обучении

Лидия Троешестова Роман Исаченко

МФТИ

6 апреля 2023 г.

Цель

Исследовать методы устранения смещения распределений p_x^- и p_x^+ без использования меток соответствующих классов.

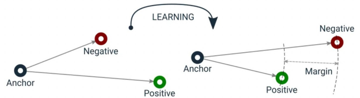
Идея

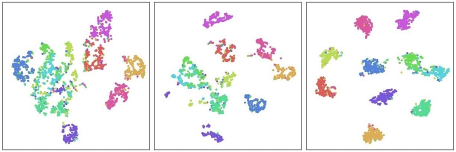
В новой функции потерь использовать оценку распределения p_x^+ в предоположении верности негативных объектов.

$$L^N_{\text{Unbiased}}(f) = \mathbb{E}_{\substack{\mathbf{x}\sim p, \mathbf{x}^+\sim p_x^+, \\ \mathbf{x}^-\sim p_x^-}} \left[-\log \frac{\exp(\text{sim}_f(\mathbf{x}, \mathbf{x}^+))}{\exp(\text{sim}_f(\mathbf{x}, \mathbf{x}^+) + \sum_{i=1}^N \exp(\text{sim}_f(\mathbf{x}, \mathbf{x}^-)} \right]$$

Постановка задачи

$p_x^+(\mathbf{x}')$ — вероятность взять $\mathbf{x}'$ как позитивный объект для $\mathbf{x}$.
$p_x^-(\mathbf{x}')$ — вероятность взять $\mathbf{x}'$ как негативный объект для $\mathbf{x}$.
τ^+ — вероятность 1 класса; $\tau^- = 1 - \tau^+$ — вероятность любого другого класса $\implies p(\mathbf{x}') = \tau^+ p_x^+(\mathbf{x}') + \tau^- p_x^-(\mathbf{x}')$





Biased | Debiased M=8 | Unbiased

Найти $L^N_{\text{DebiasedPos}}$, минимизирующее $\lim\limits_{N\to\infty} \left| L^N_{\text{DebiasedPos}}(f) - L^N_{\text{Unbiased}}(f) \right|$.

$$L^N_{\text{Unbiased}}(f) = \mathbb{E}_{\substack{\mathbf{x}\sim p, \mathbf{x}^+\sim p_x^+, \\ \mathbf{x}^-\sim p_x^-}} \left[-\log \frac{\exp(\text{sim}_f(\mathbf{x}, \mathbf{x}^+))}{\exp(\text{sim}_f(\mathbf{x}, \mathbf{x}^+) + \sum_{i=1}^N \exp(\text{sim}_f(\mathbf{x}, \mathbf{x}^-)} \right]$$

Список литературы

Ching-Yao Chuang, Joshua Robinson, Lin Yen-Chen, Antonio Torralba, Stefanie Jegelkah (2020)
Debiased Contrastive Learning

Prannay Khosla, Piotr Teterwak, Chen Wang, Aaron Sarna, Yonglong Tian, Phillip Isola, Aaron Maschinot, Ce Liu, Dilip Krishnan (2021)
Supervised Contrastive Learning

CTing Chen, Simon Kornblith, Mohammad Norouzi, Geoffrey Hinton (2020)
A Simple Framework for Contrastive Learning of Visual Representations

Sohn, Kihyuk (2016)
Improved Deep Metric Learning with Multi-class N-pair Loss Objective

Florian Schroff, Dmitry Kalenichenko, James Philbin (2015)
FaceNet: A Unified Embedding for Face Recognition and Clustering

Решение

Обозначим $h(\mathbf{x}, \widetilde{\mathbf{x}}) := e^{f(\mathbf{x})^T f(\widetilde{\mathbf{x}})}$.

Lemma

При $N \to \infty$:

$$L^N_{Unbiased}(f) \longrightarrow \mathbb{E}_{\substack{\mathbf{x} \sim p \\ \mathbf{x}^- \sim p_x^-}} \left[-\log \frac{R}{R + N\mathbb{E}_{\mathbf{x}^- \sim p_x^-} h(\mathbf{x}, \mathbf{x}^-)} \right],$$

где

$$R = \frac{1}{\tau^+} \left(\mathbb{E}_{\mathbf{x}' \sim p} h(\mathbf{x}, \mathbf{x}') - \tau^- \mathbb{E}_{\mathbf{x}^- \sim p_x^-} h(\mathbf{x}, \mathbf{x}^-) \right).$$

$$\tilde{L}^N_{\text{DebiasedPos}}(f) = \mathbb{E}_{\substack{\mathbf{x} \sim p \\ \mathbf{x}^- \sim p_x^-}} \left[-\log \frac{\mathbb{E}_{\mathbf{x}' \sim p} h(\mathbf{x}, \mathbf{x}') - \tau^- \mathbb{E}_{\mathbf{x}^- \sim p_x^-} h(\mathbf{x}, \mathbf{x}_i^-)}{\mathbb{E}_{\mathbf{x}' \sim p} h(\mathbf{x}, \mathbf{x}') + (N\tau^+ - \tau^-) \mathbb{E}_{\mathbf{x}^- \sim p_x^-} h(\mathbf{x}, \mathbf{x}_i^-)} \right]$$

Оценим неизвестные матожидания эмпирически:

$$P_{\text{emp}}(\mathbf{x}, \{\mathbf{u}_i\}_{i=1}^N, \mathbf{v}) = \frac{1}{N+2} \left(\sum_{i=1}^N h(\mathbf{x}, \mathbf{u}_i) + h(\mathbf{x}, \mathbf{v}) + h(\mathbf{x}, \mathbf{x}) \right); \; P^-_{\text{emp}}(\mathbf{x}, \{\mathbf{u}_i\}_{i=1}^N) = \frac{1}{N} \sum_{i=1}^N h(\mathbf{x}, \mathbf{u}_i).$$

$$\tilde{L}^N_{\text{DebiasedPos}}(f) = \mathbb{E}_{\substack{\mathbf{x}\sim p \\ \mathbf{x}^-\sim p_x^-}}\left[-\log \frac{\mathbb{E}_{\mathbf{x}'\sim p} h(\mathbf{x},\mathbf{x}') - \tau^- \mathbb{E}_{\mathbf{x}^-\sim p_x^-} h(\mathbf{x},\mathbf{x}_i^-)}{\mathbb{E}_{\mathbf{x}'\sim p} h(\mathbf{x},\mathbf{x}') + (N\tau^+ - \tau^-)\mathbb{E}_{\mathbf{x}^-\sim p_x^-} h(\mathbf{x},\mathbf{x}_i^-)}\right]$$

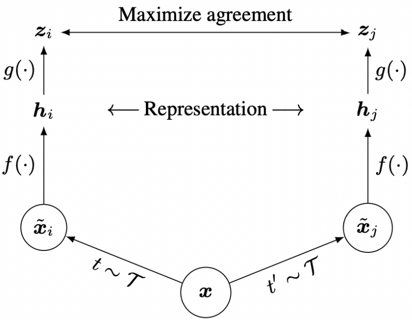
Финальная оценка:

$$L^N_{\text{DebiasedPos}}(f) = \mathbb{E}_{\substack{\mathbf{x}\sim p \\ \{\mathbf{u}_i\}_{i=1}^N\sim p_x^{-N} \\ \mathbf{v}\sim p_x^+}}\left[-\log \frac{P_{\text{emp}} - \tau^- P^-_{\text{emp}}}{P_{\text{emp}} + (N\tau^+ - \tau^-)P^-_{\text{emp}}}\right].$$

Theorem

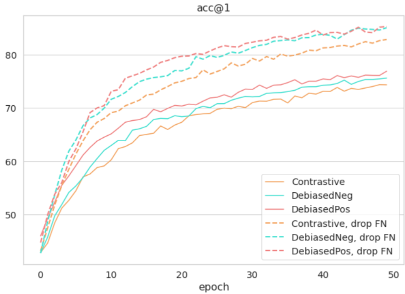
Для произвольного эмбеддинга f и произвольного $\delta > 0$ существует достаточно большое N, что

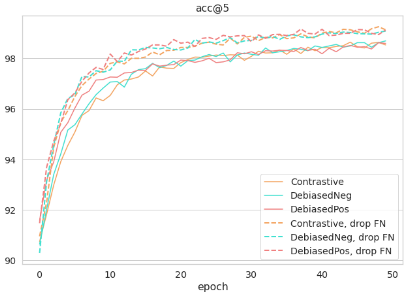
$$|\tilde{L}^N_{DebiasedPos}(f) - L^N_{DebiasedPos}(f)| \leq \left[\left(1 + \frac{\tau^-}{\tau^+} + \delta\right)\sqrt{\frac{\pi}{2N}} + \left(1 + \frac{1}{\tau^+}\right)\sqrt{\frac{\pi}{2N+2}}\right] e^{3/2}$$

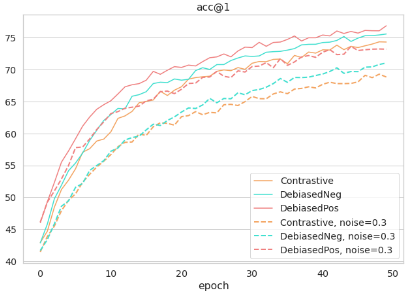


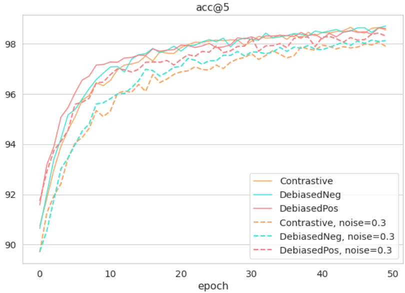
- $\mathcal{T}$ – семейство аугментаций (color distortion, gaussian blur)
- Семплируются 2 аугментации $t, t' \sim \mathcal{T}$, применяются к каждому объекту.
- Обучаем сеть-энкодер $f(\cdot)$ и MLP сеть-проекцию $g(\cdot)$, максимизируя соответсвтие представлений.

Эксперимент с удалением FN и с добавлением FP

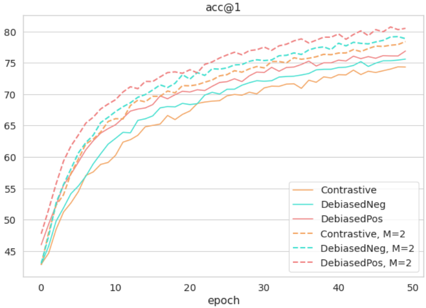


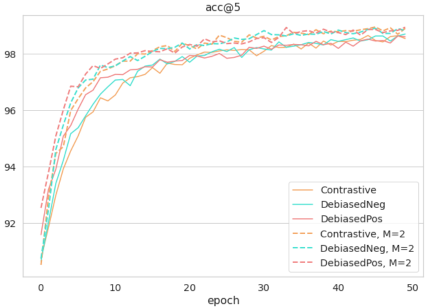






Эксперимент с увеличением кол-ва положительных семплов





Заключение

- Лосс с устраненным FP-смещением добивается точности, сопоставимой с DebiasedNeg, и имеет превосходство на ранних эпохах.
- В случае удаления ошибок FN функции потерь DebiasedPos и DebiasedNeg имеют сопоставимые точности.
- При сильном зашумлении датасета, когда FP-rate увеличен, DebiasedPos имеет большое преимущество в точности над DebiasedNeg.
- При увеличенном кол-ве положитлеьных семплов точность для всех рассмотренных функций потерь возрастает.