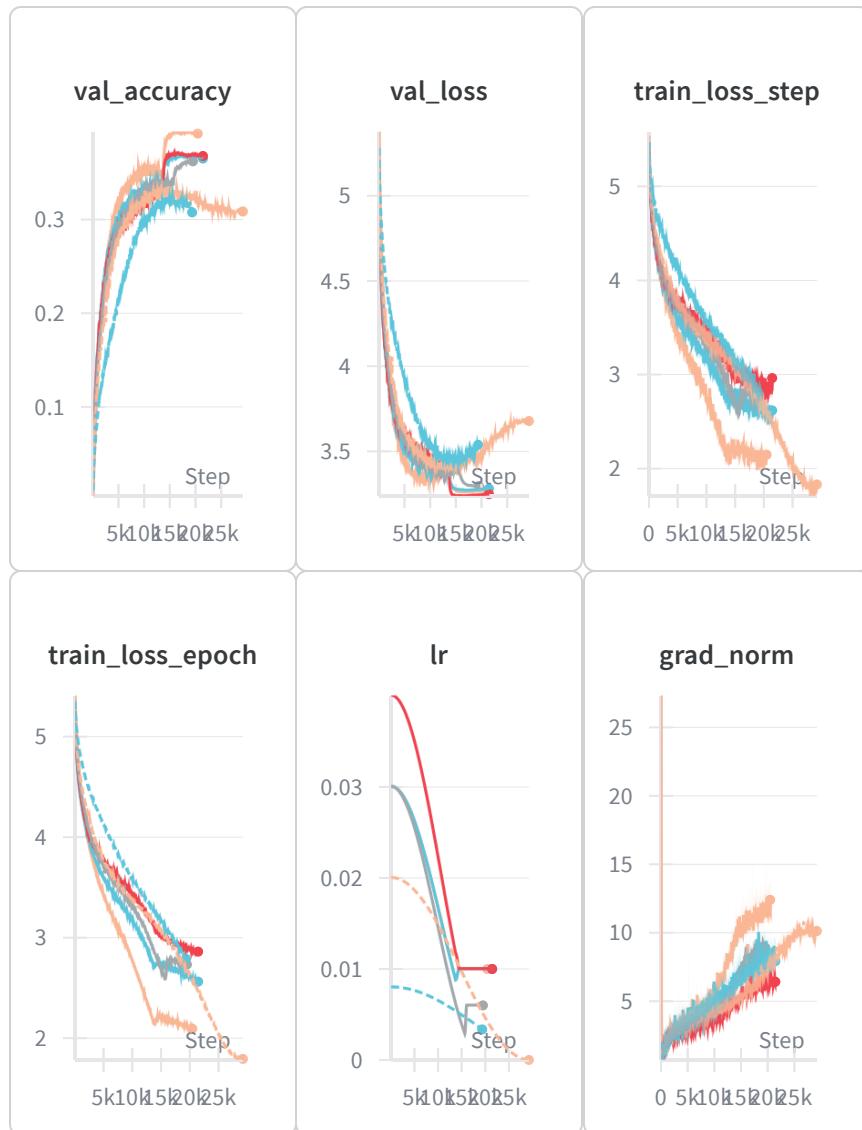


BDZ-1

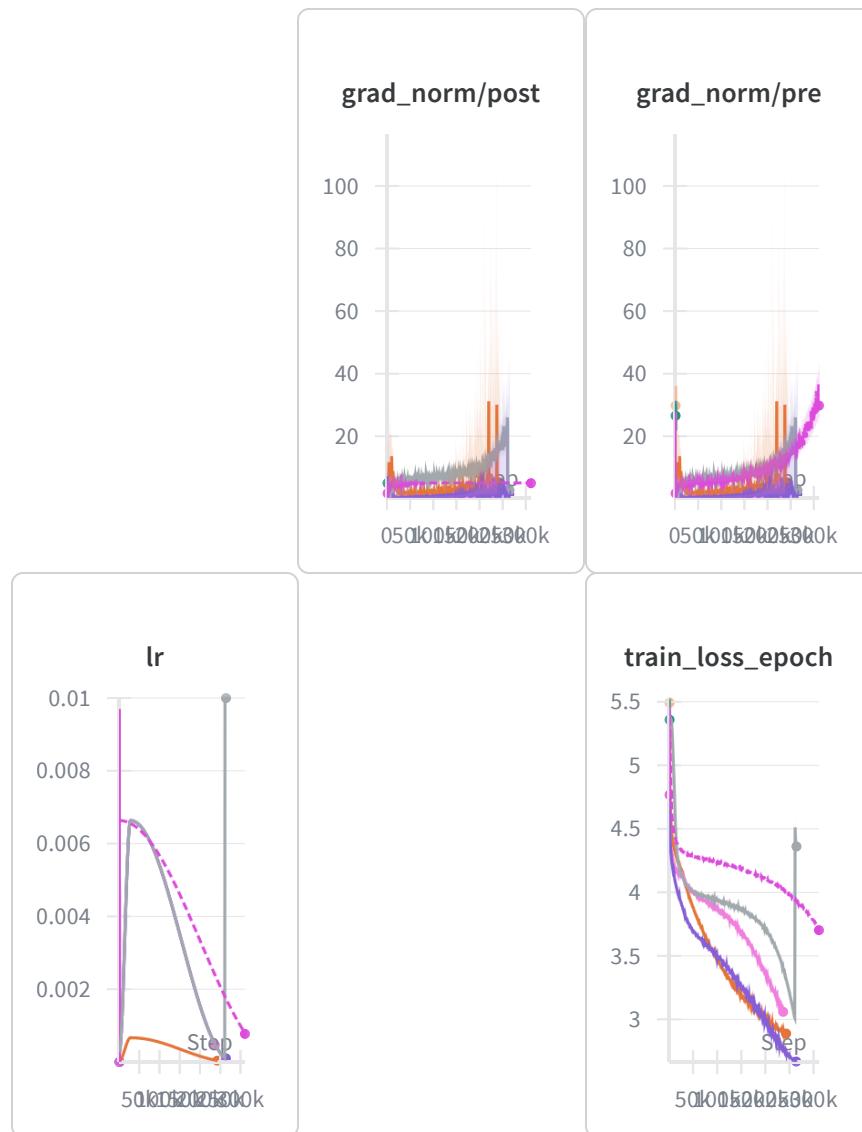
Nikita

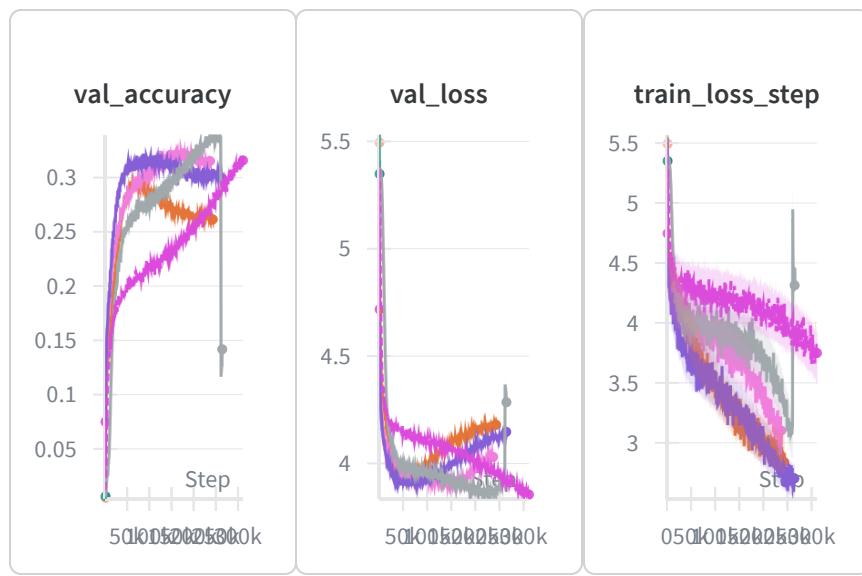
Created on January 22 | Last edited on January 22



Если кратко для обучения использовался Resnet18 с аугментациями Resize, RandomCrop и RandomAugs. Хорошо себя показал оптимизатор SGD так как более стабильный, Adam быстро поднимается но потом уходит в горизонталь.(сверху будет лучше видно) Потом начал обучать Resnet50. Было очень много ablation и

для той и для той модель(использовалась optuna). Примеры некоторых сверху. На данном этапе я еще начал клиппать градиенты, но на графике показан градиент до клиппинга(по крайне мере для лучшей модели я точно клипал). Изначально планировалось что это будет моим бейзлайном, но идея с ArcFace провалилась. Ниже результаты, которые смог получить. Также использовал SWA в конце обучения что видно по графикам.





Тут я уже начал логировать градиенты до и после.

resnet_18_arcface_m_60 и resnet_18_arcface_m_10 выглядят многообещающе, правда для resnet_18_arcface_m_10 val_loss начинает подниматься хоть это не о чем не говорит(так как обязательно на валидации тоже классам очень хорошо разделяться). Для ArcFace я написал свой даталоадер чтобы он разделял разные классы каждый раз поэтому batch_size маленький(-200). Возможно есть способы его увеличить когда классов мало, но я их не нашел. Еще как будто swa не помогает, хотя для resnet_18_arcface_m_10 я установил его слишком высоким. Думаю это идея добивается и результат можно хороший получить но надо долго мучится с гиперпараметрами и надо батч сайз увеличить какимто образом. Еще ауги может правильные нужны.(типа поворота). Ну из того что я читал ArcFace идея уже скорее используется для файнтюона

Created with ❤️ on Weights & Biases.

https://wandb.ai/solonicynnikita6-yandex/DL-BDZ-1_exp/reports/BDZ-1--VmlldzoxNTcxMjI4Mw