

[Share](#)[Comment](#)[Star](#)

...

CNN-report

[Gleb Dolgushev](#)

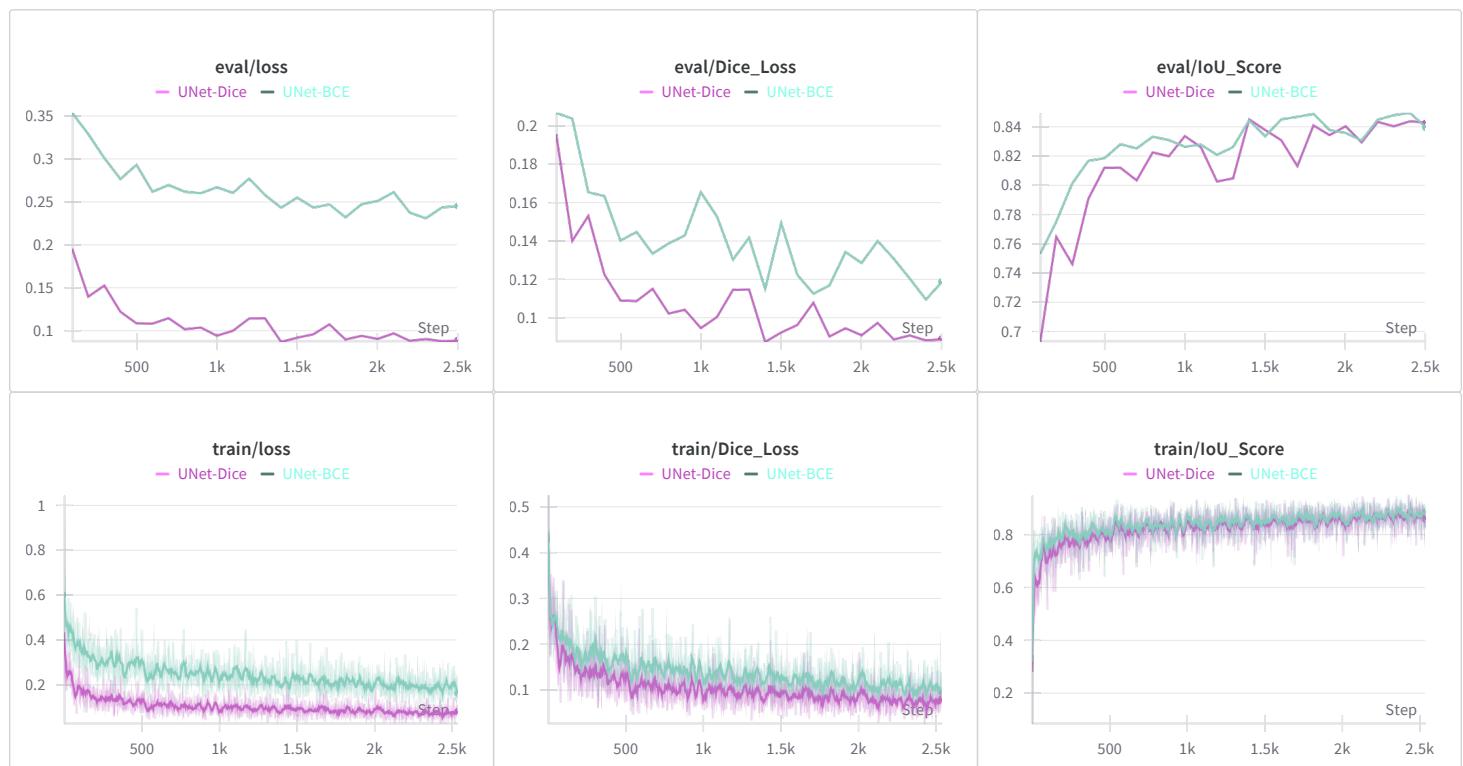
Created on March 14 | Last edited on March 15

1. Logging demonstration.

Покажем корректную работу логирования на примере UNet-BCE (с функцией потерь кросс-энтропия) и UNet-Dice (с функцией потерь DiceLoss). Логирование гиперпараметров и архитектуры модели доступно в самом интерфейсе wandb.

1.1 Evaluation / Train quality and losses

Пункты 1 и 2 задания



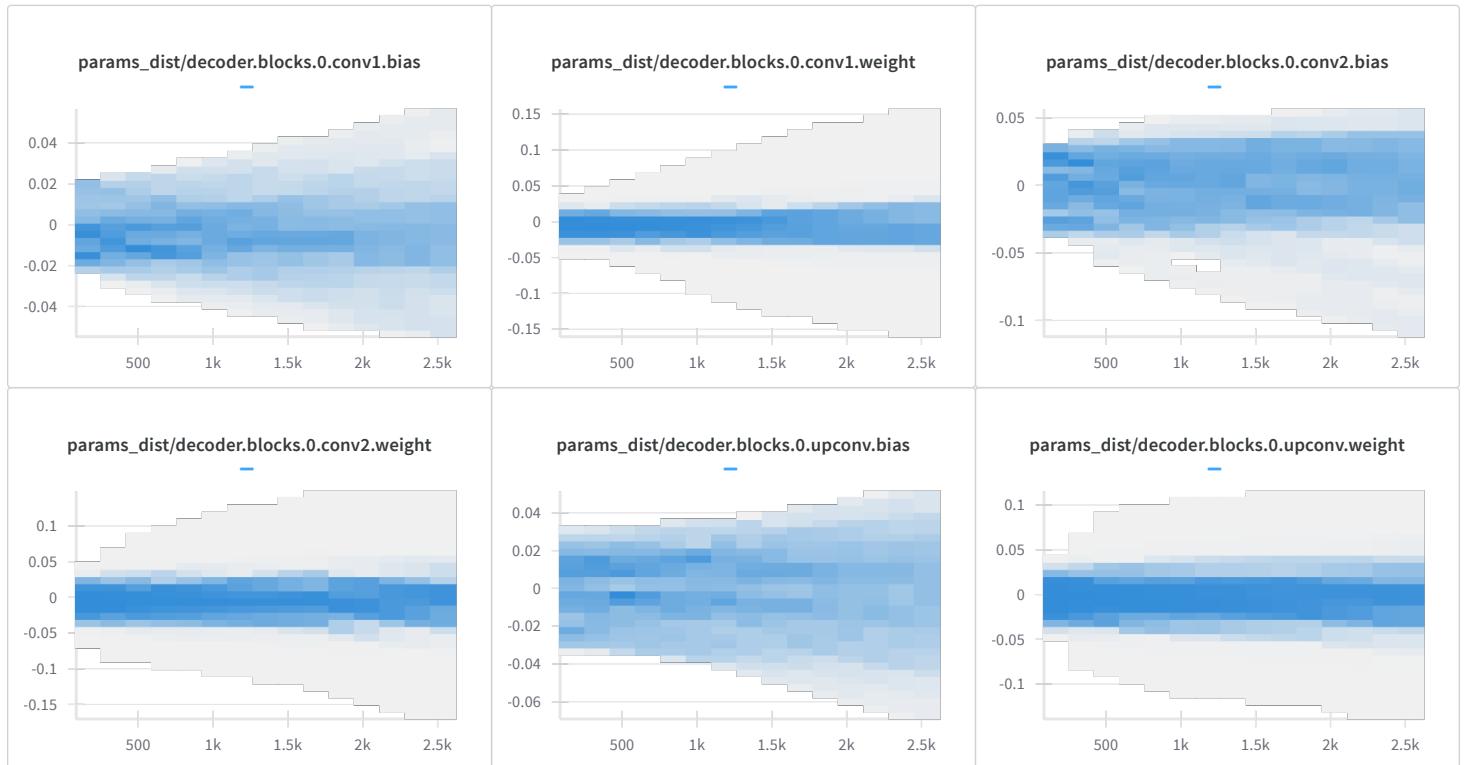
1.2 Evaluation / Train visualization

Пункт 3 задания



1.3 Weights distribution

Пункт 6 задания (Отдельно для UNetBCE)



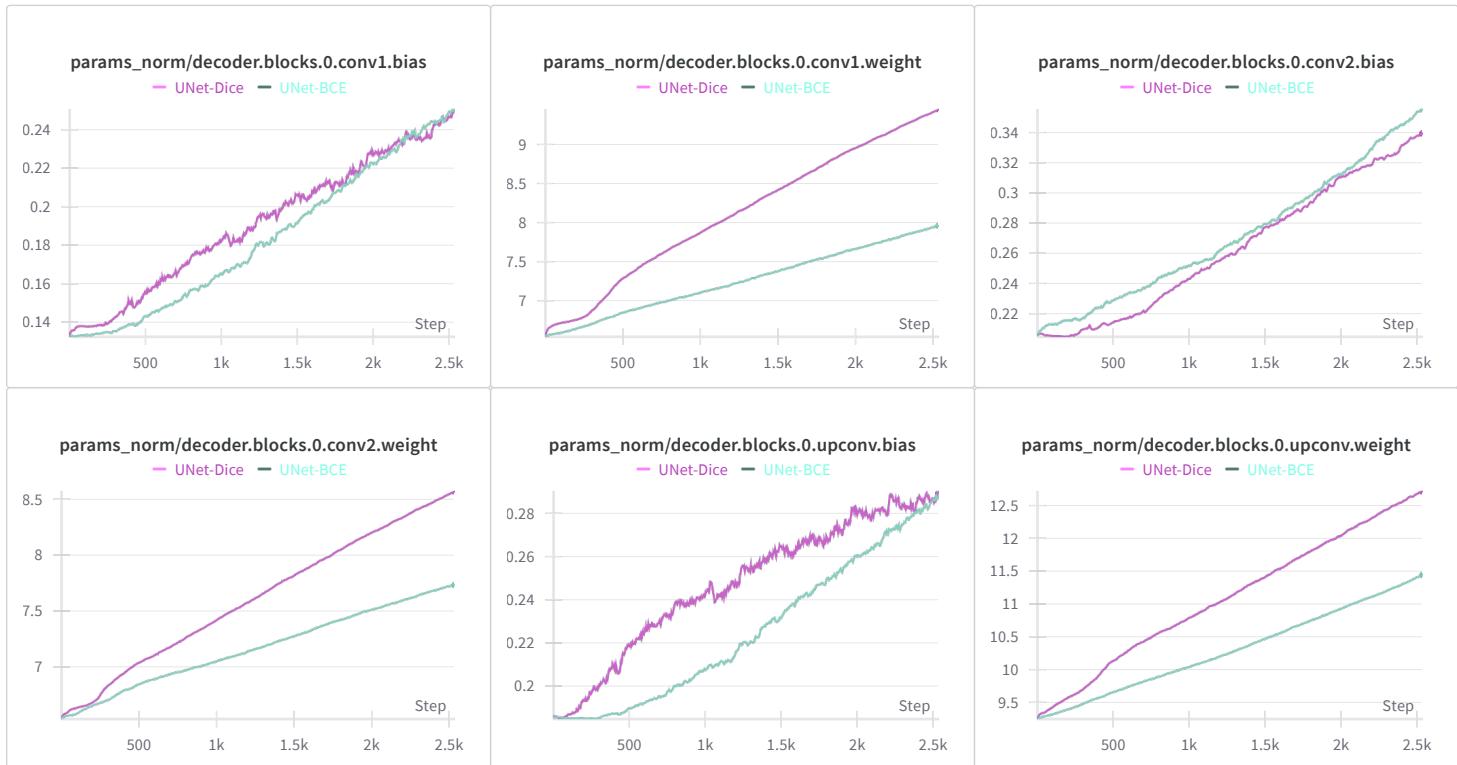
1.4 Gradients distribution

Пункт 6 задания (Отдельно для UNetBCE)



1.5 Weights norm

Пункт 7 задания



1.6 Gradients norm

Пункт 7 задания



1.7 Activations as images

Пункт 8 задания. В следующих двух пунктах логировались только первые 9 каналов слоя активации, так как логирование

всех привело бы к нагромождению результатов.



Отметим, что многие активации хорошо интерпретируемы, например, на слое 0 (9-ый канал) выделяется губы, на слое 1 - лица (3-й канал).

1.8 Gradients as images

Пункт 9 задания

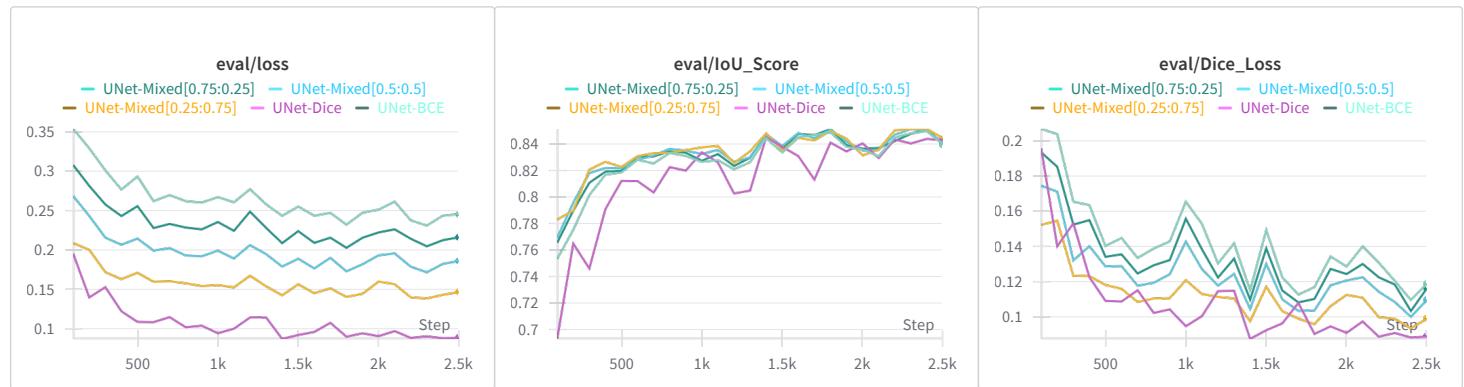


2. Comparison of properties of UNet for different loss functions

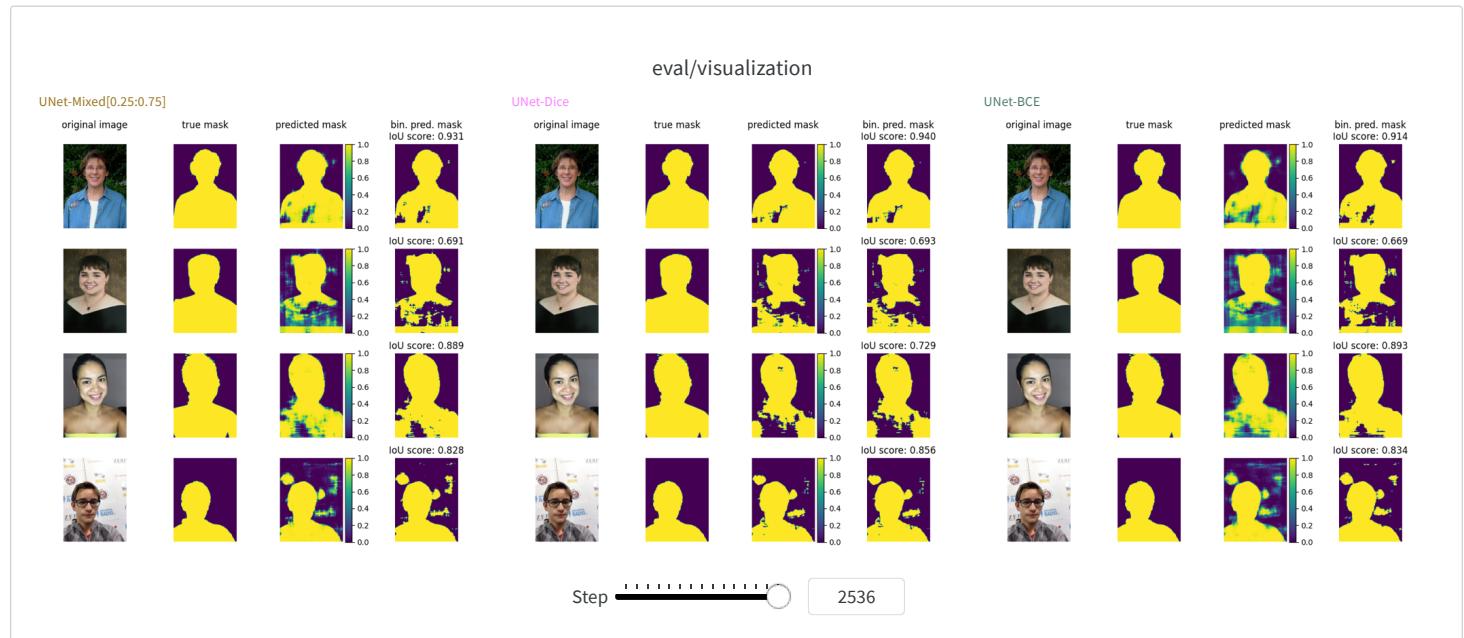
Проанализировать особенностей использования как функции потерь кросс энтропии или DiceLoss можно оценить по предоставленным выше графикам. Сделаем следующие выводы:

1. Лучшего качества позволяет достичь обучение с функцией потерь BCE
2. Выходы модели, обученной на оптимизации BCE более интерпретируемы, в то время как модель, обученная с DiceLoss похожа на излишне уверенный классификатор (все вероятности либо около 0, либо около 1)
3. При использовании BCE модель обучается за меньшее число итераций

Далее добавим в рассмотрение эксперименты, в которых как функция потерь бралась линейная комбинация BCE и DiceLoss (Коэффициенты указаны в названии моделей в соотношении [коэффициент BCE:коэффициент DiceLoss]).

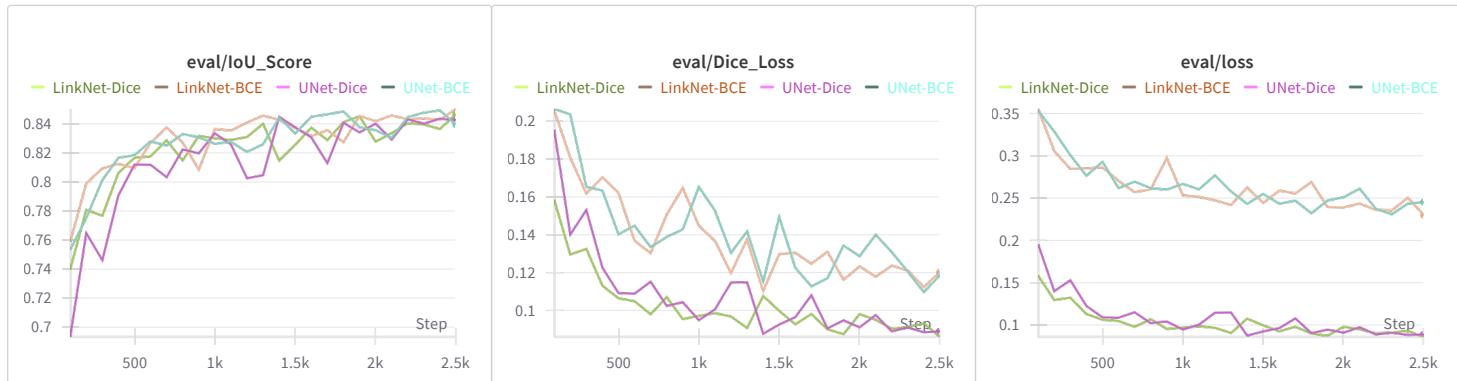


Как можно видеть, смешивание функций потерь помогает и улучшить качество (лучший результат достигнут при соотношении [0.25:0.75]), и оставить вероятностную интерпретацию предсказаний (даже при довольно небольшом коэффициенте при BCE на границах предсказания модели уже не такие уверенные)

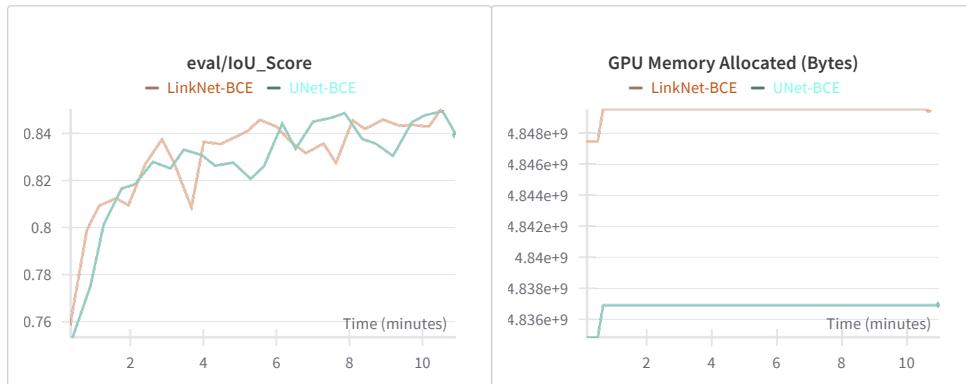


3. Link-Net

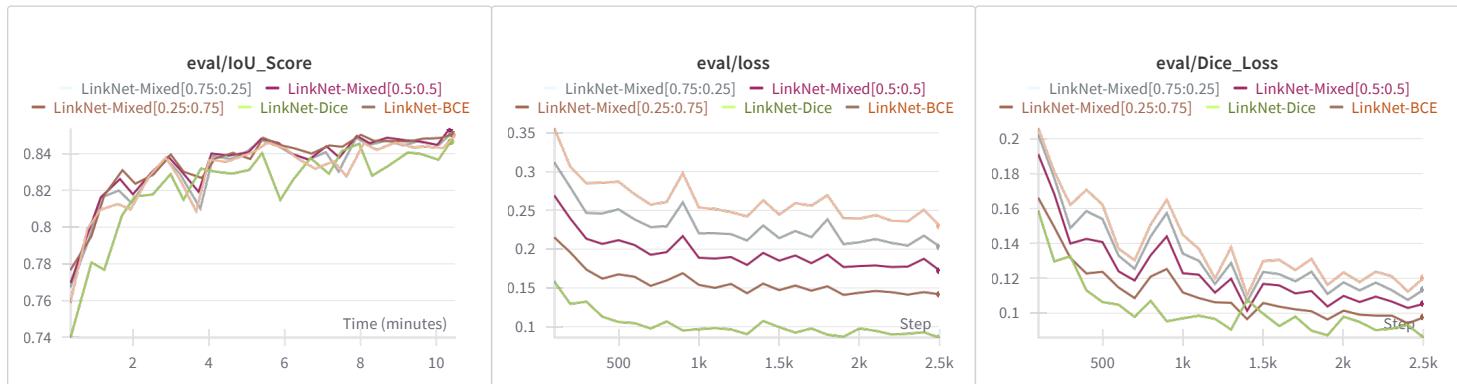
Сравним Link-Net с BCE и DiceLoss с UNet с аналогичными функциями потерь. Наблюдаем ту же тенденцию с превосходством моделей, обученных с использованием BCE. При этом Link-Net показывает лучшее качество.



Интересно также сравнить время работы моделей (Видим небольшое преимущество у Link-Net). При этом новая модель требовала больше памяти во время обучения



Аналогично UNet проведем перебор различных комбинаций BCE и DiceLoss. В данной ситуации лучше оказывается соотношение (0.5:0.5).



Created with ❤️ on Weights & Biases.

<https://wandb.ai/kellesett-msu/CNN-workshop/reports/CNN-report--VmldzoxMTc5Mjk2Mg>

