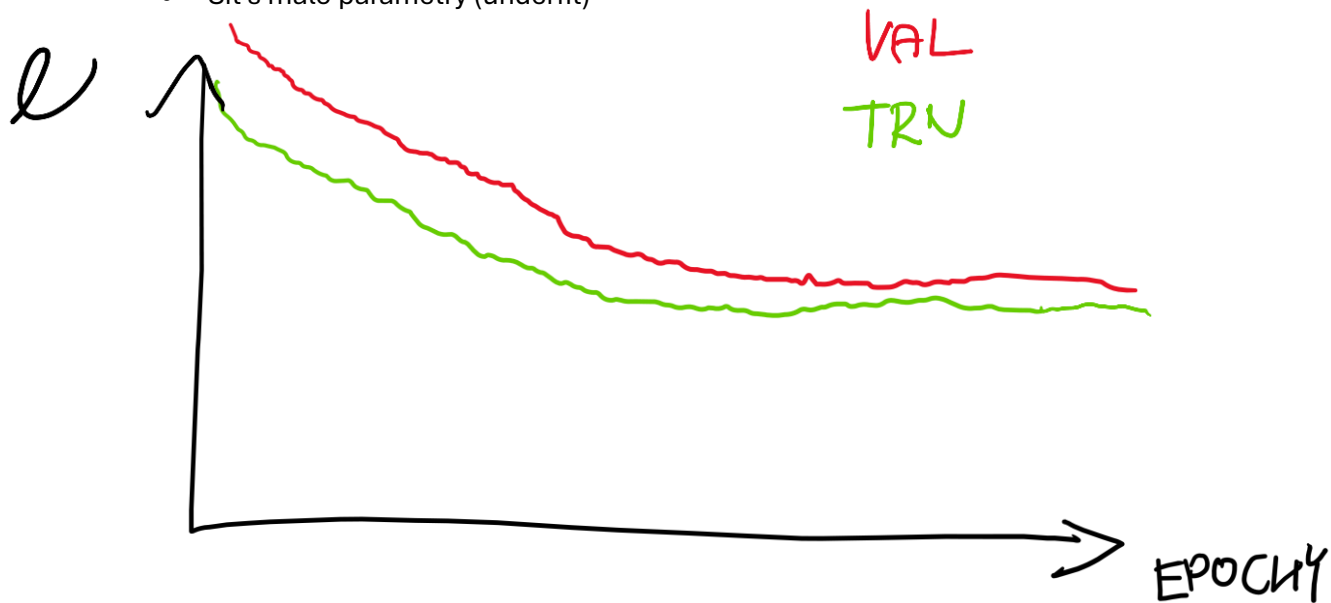
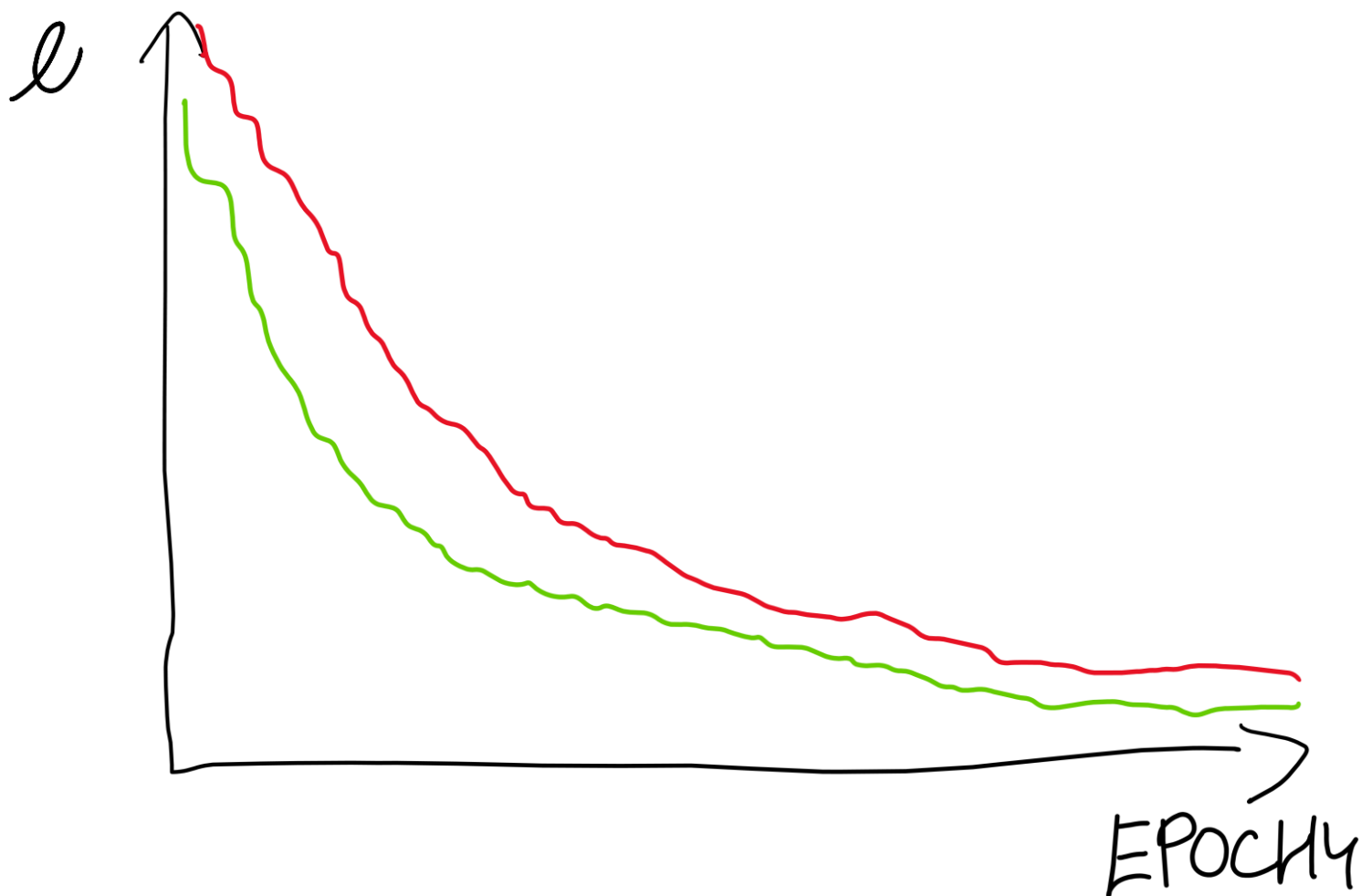


1. Grafy

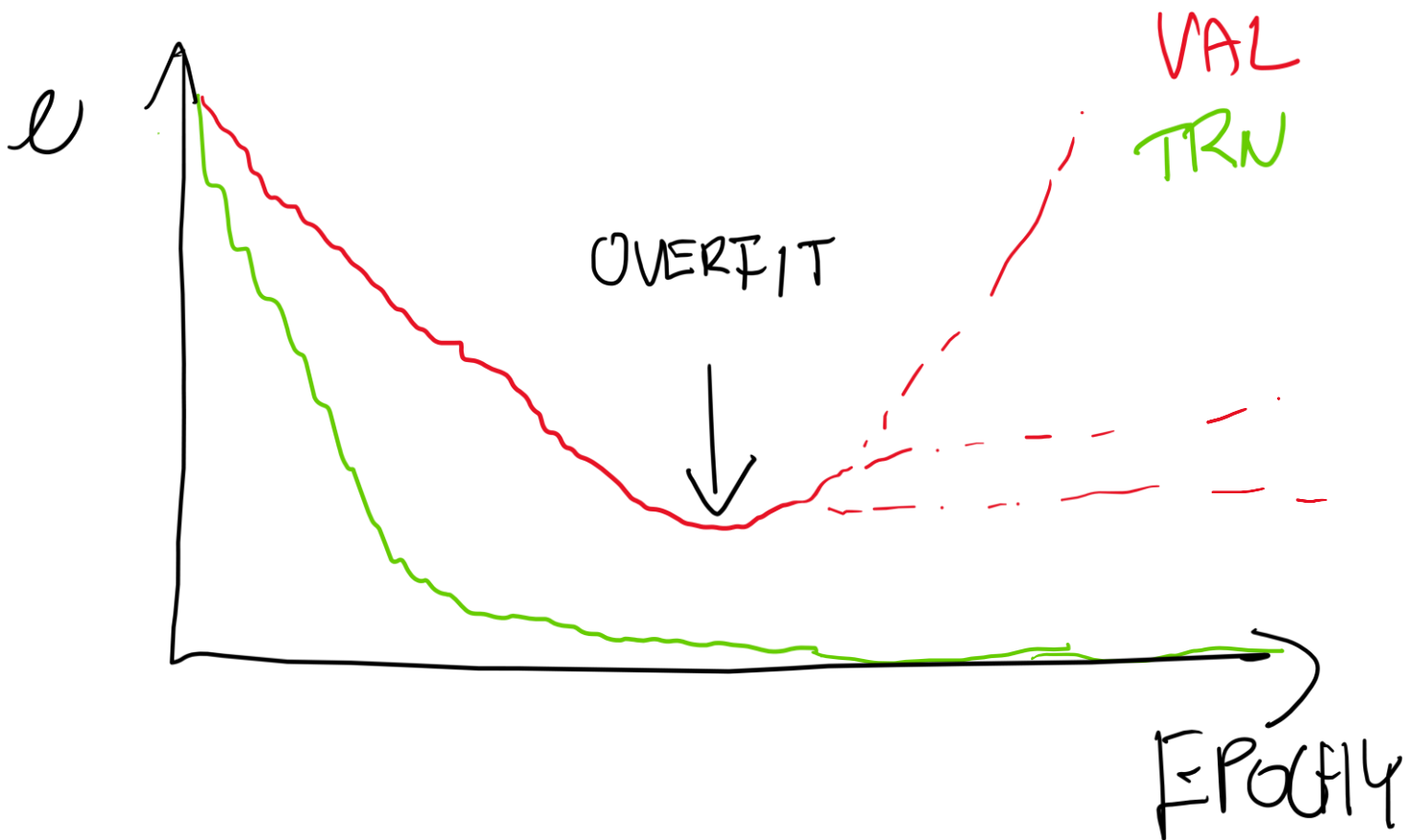
- Síť s málo parametry (underfit)



- Ideální stav



- Příliš mnoho parametrů (overfit)



2. Mix up technika

Lineární kombinace vstupů očekává lineární kombinaci labelů na výstupu.

Výhody:

- Obrovské množství nových dat
- Kombinace zdravé a nemocné plíce může udělat "polozdravou" plíci, tedy plíci, kde něco začíná
- Potenciálně lepší výsledky

Nevýhody

- Potřeba dobrého alignmentu, aby někde nevznikalo např. více žebírek
- Ty snímky nejsou reálné
- Mohou vznikat artefakty (např. lehký náznak kardiostimulátoru)

Technika se dá použít v závislosti na našem cíli. Pokud bychom ji chtěli použít jen jako hint pro lékaře (tak jak to aktuálně je), můžeme chtít nejlepší výsledky i za cenu toho, že použijeme neexistující snímky a lékař pak sám rozhodne, jestli jsme našli správný nebo špatný výsledek.

Pokud by ovšem takováto kontrola nenastávala a my rozhodli špatně, mohlo by se zkoumat, proč tato chyba nastala a při zjištění, že jsme používali takovou augmentaci, by mohl nastat problém. Pro tento scénář by tedy stačily pouze rigidní transformace, které mohou v reálném světě nastat.

3. Mamografický snímek, která NN?

Určitě musí jít o konvoluční neuronovou síť, ale jestli se jedná o U-NET, ResNet se podle mě nedá odhadnout.

4. Jak vytvořit syntetická data?

Mohli bychom použít GAN, kamarád jako semestrální projekt používal GAN aby generoval snímky magnetické resonance mozku a mělo to velmi dobré výsledky.

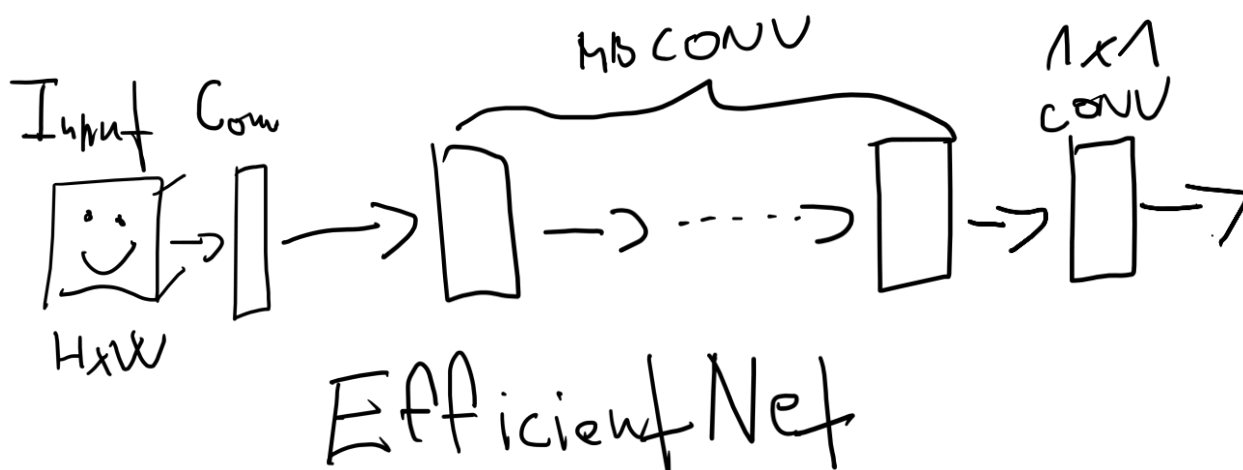
5. MLP trénink

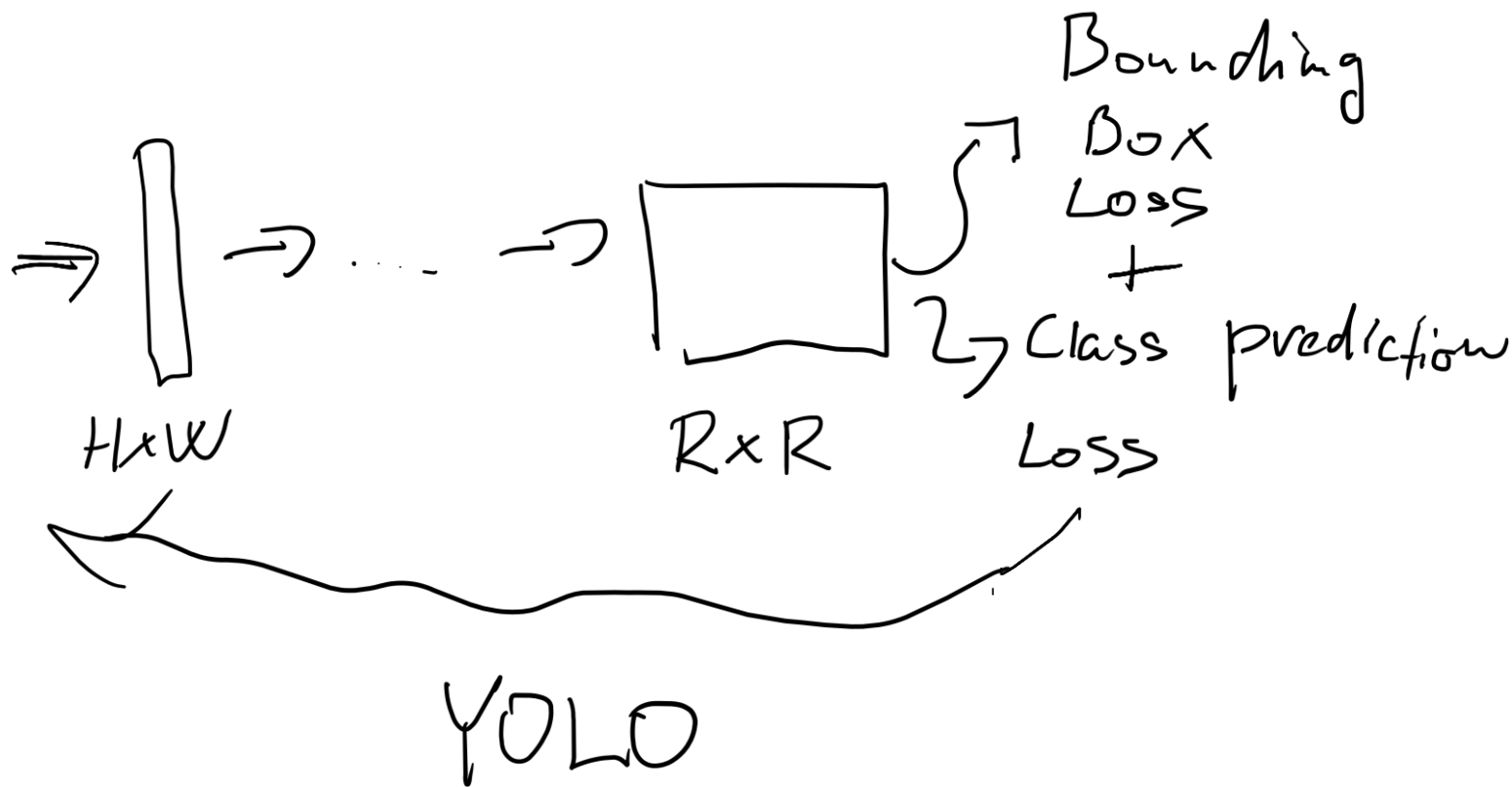
- A. Pravděpodobně nejjednodušší, tedy RELu a BatchNorm
- B. Může se buď použít neuronová síť s menším počtem parametrů, aby neoverfitovala. Nebo ukládat během tréninku po každé epoše síť, která měla doposud nejmenší validační chybu. K tomu mu ještě případně přidat augmentaci dat.
- C. Accuracy je 40 %. My používali pro tento případ triplet loss funkci, kde jsme získávali hard negatives a hard positives.

6. Zaostření obrázku

Pokud máme více obrázků stejné věci, můžeme použít průměrování, což snižuje šum. Můžeme použít nějaké digitální filtry na ostření hran, což může a nemusí pomoci.

7. EfficientNet & YOLO





8. Batch size

Pokud se vejde batch size 20 na GPU není asi důvod snižovat na 5, protože to bude pomalejší a více stochastické. Nicméně snížení batchsize na čtvrtinu by znamenalo, že se mi do sítě vejdou 2x větší obrázky.