- describe your data processing pipeline and comment on your decisions

* finální model trénován na datasetu s 20k páry (40000), rozdělený na train (35000), val (4000), test (1000)
* evaluace konfigurace sítě na menších modelech s 5k páry
* používáme generátory pro čtení fotek z disku, preprocessing
* vstupem do neuronové sítě jsou dvě matice fotek, při trénování label 1/0 zda jsou stejné nebo ne
* cachování předzpracovaných obrázků - nefungovalo načítání (pickle encoding chyba? zkoušeno více přístupů, ani jeden příliš dobře nefungoval), mohlo proces trénování zrychlit

- describe your model, process of hyperparameters selections, why you pick the chosen architecture

* nejlepší performance ze zkoušených architektur, dosáhla přinejlepším 0.82 accuracy, bez zvyšování val\_loss
* zkoušel jsem síť redukovat, zvětšovat
* proč ne grid search hyperparametrů - tensorflow mi neuvolňuje nikdy vram, spadlo mi to

- describe training process (what were values of validation loss and training loss, if you encountered overfiting during training, how did you deal with that, etc…)

* zprvu se model příliš neučil, nejvíce znatelnou změnou byla náhrada feature extractoru na m-r50x z tensorflow hub, zkoušel jsem i jiné feature extractory (VGGFace, MobileNet… - výsledky horší nebo identické, mobilenet výrazně rychlejší), trénovat resnet od začátku, contrastive loss (zde mi accuracy šla k nule - nevím jestli byla chyba u mě…)
* overfitting - ano, řešeno přidáním dropout vrstvy, L2 regularizer
* postupně snižována komplexita modelu
* konfigurace sítě
  + zvolená konfigurace má nejvyšší úspěšnost, 0.81 accuracy na redukovaném datasetu 5k párů, na finálním datasetu
  + různá doporučení, ale tohle mělo nejlepší performance (accuracy)
    - doporučení dát batch norm před aktivační funkci
      * nejprve dropout, pak batch norm - accuracy neroste, pohybuje se kolem 0.4ne
    - batch norm po feature extractoru - val\_accuracy 0.79
    - bez batch norm výrazně horší výsledek, pod 0.7
  + počet parametrů - při testování konfigurací sítě zůstával stejný, ale redukcí se povedlo zrychlit trénování a zvýšit accuracy, model byl příliš komplexní (vysoký počet parametrů na malou trénovací sadu)