Below are details for each of the training iterations:

**1:** Fully functional network, first definition. Training works for both parts without memory leaks. Weights exported after both training sessions.

* Text batch size: 128
* Img batch size: 128
* FCN: lr =1e-4
* CNN: lr =1e-5

**2:** Same as above but with the following changes:

* Text batch size: 512
* FCN: lr = 1e-5

**3:** Same as 2 but with pruned, smaller dataset

**4:** Same as 3 but with 8 epochs of training