[*Mnih et al.*, 2015] report using the Deep Q-Network to let the computer to learn how to play the Atari 2600 games. This paper has the detailed information about the structure, parameter and the way they train the algorithm. It has two key ideas, (1) use the experience replay, and iterative update that adjusts the action-values towards target values. It uses the discounted award function, which discount the future actions for the rewards. The learning essentially is a regression problem to estimate the Q-value from the observed images. The code is written all in Lua, but there is a similar project using it to train the flappyBird, which you can find it here: <https://yanpanlau.github.io/2016/07/10/FlappyBird-Keras.html>.

[*Araya-Polo et al.*, 2018] talks about their way to do tomography velocity model using deep neural network. They are still only using the synthetic data by first randomly generating velocity model and then generate the seismic waveforms, and use these as the pair to train the network. After that, they will generate more velocity model and test the model against these newly generated models. Also note that, their output needs to be processed in order to have the same demission as the velocity models.