The need to preserve and exchange written information is central to the human society, with handwriting satisfying such need for several past millenia. Unlike optical character recognition of typeset fonts, which has been throughly studied in the last few decades, the task of handwritten text recognition, being considerably harder, lacks such attention. In this work, we study the capabilities of deep convolutional and recurrent neural networks to solve handwritten text extraction. To mitigate the need for large quantity of real ground truth data, we propose a suitable synthetic data generator for model pre-training, and carry out extensive set of experiments to devise a self-training strategy to adapt the model to unnanotated real handwritten letterings. The proposed approach is compared to supervised approaches and state-of-the-art results on both established and novel datasets, achieving satisfactory performance.