Abstract：

Many localized languages struggle to reap the benefits of recent advancements in character recognition systems due to the lack of substantial amount of labeled training data. This is due to the difficulty in generating large amounts of labeled data for such languages and inability of deep learning techniques to properly learn from small number of training samples. We solve this problem by introducing a technique of generating new training samples from the existing samples, with realistic augmentations which reflect actual variations that are present in human hand writing, by adding random controlled noise to their corresponding instantiation parameters. Our results with a mere 200 training samples per class surpass existing character recognition results in the EMNIST-letter dataset while achieving the existing results in the three datasets: EMNIST-balanced, EMNIST-digits, and MNIST. We also develop a strategy to effectively use a combination of loss functions to improve reconstructions. Our system is useful in character recognition for localized languages that lack much labeled training data and even in other related more general contexts such as object recognition.

主要内容：很多人想在本地语言上做文字识别，因为数据集不够很难以完成。作者发明一种方法通过在小样本中加入噪声，以扩充样本量的方法来提高识别率。分别在EMNIST-balanced, EMNIST-digits, and MNIST.进行了测试，效果很好。

Introduction

暂时略过，后期需要再进行补充。

第三节 方法

提出胶囊网络和一个解码器网络

胶囊网络结构：

卷积1 ：64个3\*3核 stride 2

卷积2 ：128个3\*3核 stride 1

卷积3 ：256个3\*3核 stride 2

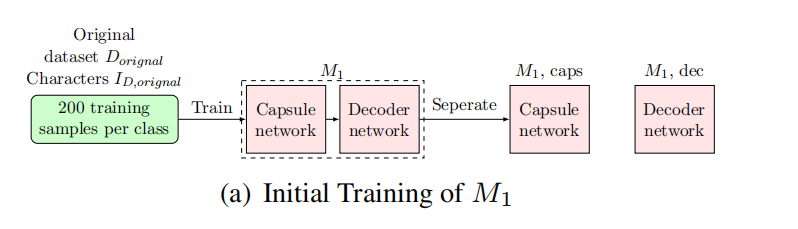
胶囊层 ：32通道 8维胶囊 每个胶囊包含8个卷积单元 9\*9的核 stride 2

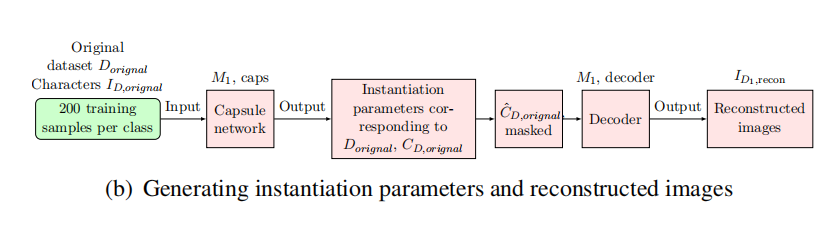
特征胶囊层：全连接层，16维 分出M个胶囊与类别对应

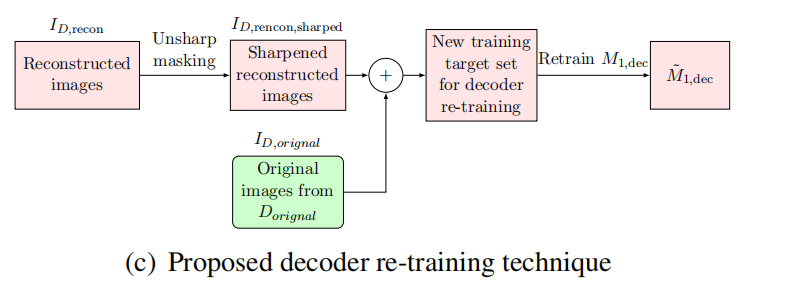
解码器网络：

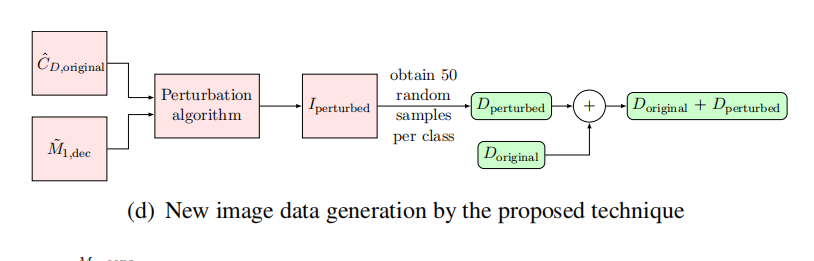
5个反卷积层

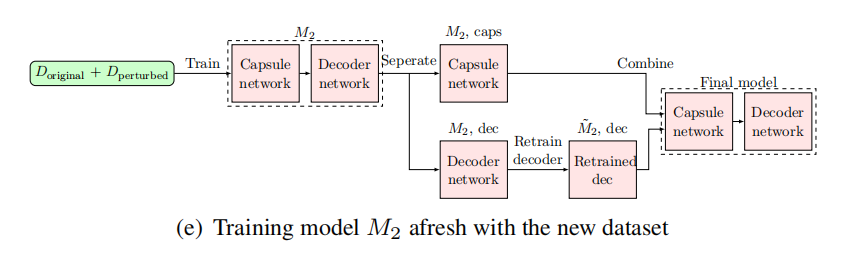
输出28\*28











结果中分为两部分，首先对识别结果进行了分析。然后对生成的图片进行了分析，并于CGAN做了比较，更优更实用。