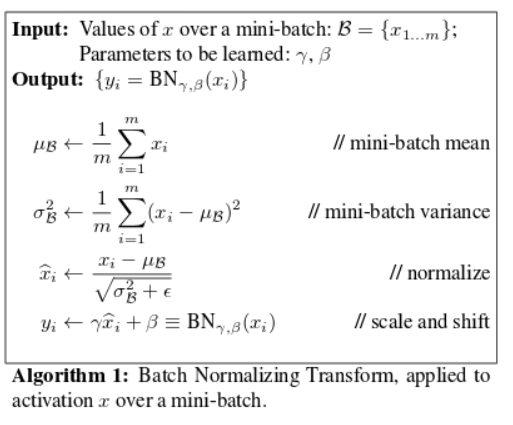
To increase the stability of a neural network, batch normalization normalizes the output of a previous activation layer by subtracting the batch mean and dividing by the batch standard deviation.



А уже параметры для скейл и шифт подбирает нейросеть