# [谈谈如何训练一个性能不错的深度神经网络](http://blog.csdn.net/kuaitoukid/article/details/45821891)

分类： [深度学习](http://blog.csdn.net/kuaitoukid/article/category/2434733) [机器学习](http://blog.csdn.net/kuaitoukid/article/category/2434731) [caffe](http://blog.csdn.net/kuaitoukid/article/category/2654799) [CNN](http://blog.csdn.net/kuaitoukid/article/category/2565477) [心得](http://blog.csdn.net/kuaitoukid/article/category/1273252)2015-05-18 16:25 649人阅读 [评论](http://blog.csdn.net/kuaitoukid/article/details/45821891#comments)(6) [收藏](javascript:void(0);) [举报](http://blog.csdn.net/kuaitoukid/article/details/45821891#report)

深度学习大火，将各个数据集的state of the art不断地刷新，到了开源代码一放出，有种全民皆可刷排名的节奏。

不过可别把刷数据想的那么简单，不然大家去哪发paper，怎么混饭吃= = 但是我不想发paper就想占坑刷数据怎么办，看到cifar10都尼玛刷到了95%了，我这用caffe自带的小demo才得出78%的结果，caffe你确定不是在骗我？

caffe确实没在骗你= =今天我给大家介绍一下如何刷出一个性能接近paper的神经网络

以CNN为例，基本分为三步：

第一步使用leaky relu，dropout（具体参见[blog.kaggle.com/2015/01/02/cifar-10-competition-winners-interviews-with-dr-ben-graham-phil-culliton-zygmunt-zajac/](http://blog.csdn.net/kuaitoukid/article/details/blog.kaggle.com/2015/01/02/cifar-10-competition-winners-interviews-with-dr-ben-graham-phil-culliton-zygmunt-zajac/)）

第二步，数据扰动，将数据上下左右平移，放大缩小，泛绿，泛红，反色等等，做很多的合理的扰动，

第三步，固定步长学习，直到训练不动，找一个高精度的solverstate作为起点，将学习率下降再训练，按理说降到1e-4就训练的差不多了

其实当你研究多了就发现，真正提高性能的还是第二步，其他的只能说是锦上添花，数据扰动才是根本，当然这也揭示了分类器本身的缺陷。

当然了，有人问，你网络结构还没给呢，这个嘛，paper跟实验接触多了，自己自然就会设计了，我倒是觉得网络结构不是主要的，因为CNN的致命缺陷其他分类器也都有，要解决只能说是都一起解决。

MNIST我靠数据扰动将结果刷到了99.58%，结构很简单粗暴无脑，cifar10扰动太少只是88%，做多了上90%应该很轻松，ImageNet，呵呵，看看金连文老师微博上对百度在ImageNet上的评论你就知道我想说什么了。

（百度把指标刷到了4.58%，主要的工作就是（1）更多的银子(144块GPU集群)（2）更大的网络（6个16层212M的模型集成）（3）更多的数据（每张图合成出上万种变化）——金连文）