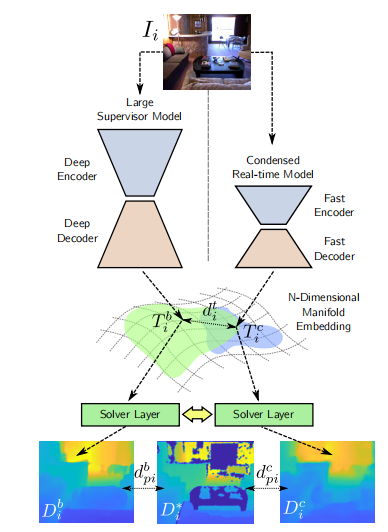
论文阅读笔记：

CReaM: Condensed Real-time Models for Depth Prediction using Convolutional Neural Networks（2019）

**内容**：提出了一种新颖的实时结构预测框架，可以在NVIDIA-TX2上预测每秒30帧的深度。在资源受限的移动平台上将实时机器人和机器学习相结合。更具体地说，提出了第一项工作，它执行单一图像深度预测。

（在这篇论文中，第一它提出了一种实时结构预测框架。它的新颖点在于它在移动平台上实现了实时机器人和机器学习的结合。一般更多层更广泛的模型可以带来更好的结果，但是它无法部署在资源受限的平台上，论文中提出的不仅可以在实现并且还优于更密集的架构。）

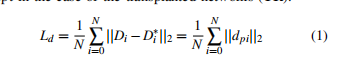
**模型架构：**作者这篇论文仅仅是完成了这个项目的第一步实现了在移动平台的运行的能力。



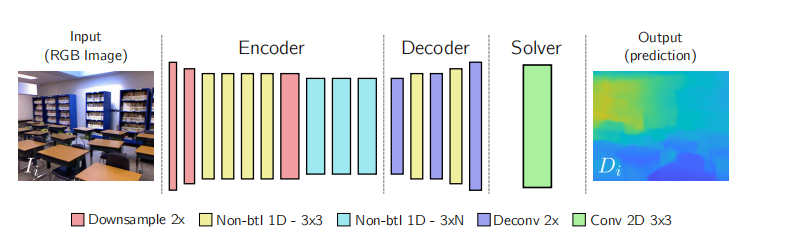
这是他们train用来执行模型压缩的方案，引入了张量损失（由网络作为输入到solver layer而产生的中间激活之间的差异，主要模仿潜在空间或者嵌入较大模型的倒数第二层）图是两种方案的对比，单独transfer的方案更加有效。

1. 随机模型

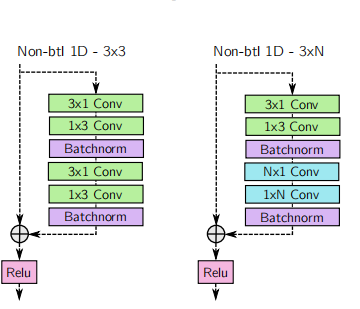
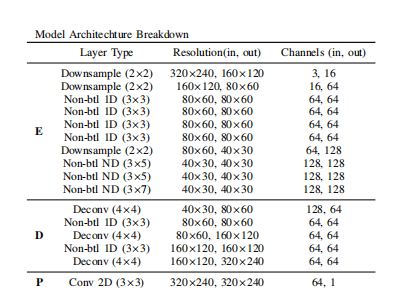
选择预测和实际的L2的距离作为回归时的损失函数，主要训练随机初始化模型



2、压缩网络（深度网络+Tensor Supervied Network）



编码层+解码层+预测层（监控网络移植）



Non-btl 主要指通过该层保持不变的通道数 N\*1,1\*N，3\*1，1\*3实行非对称卷积。

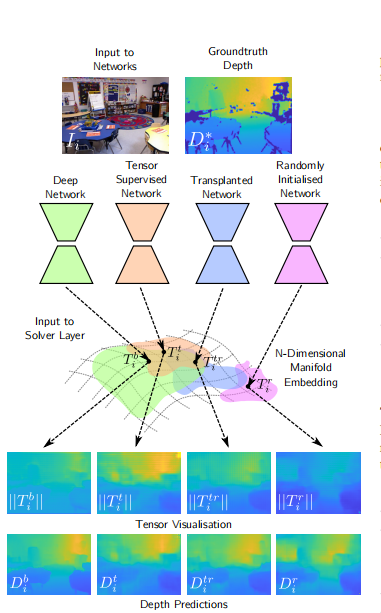
两个3\*5和一个3\*7为了增加感受野 （感受野不怎么理解，增加有什么作用）

最后一层可以transfer到压缩网络上。

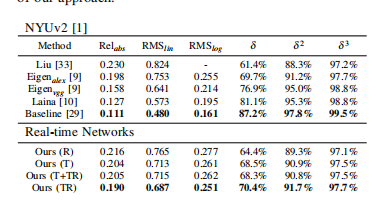
1810bfa9db8e6eb48b886f604996c4b张量损失（模仿深层模型倒数第二层的激活，Ti对应激活的张量）训练直到收敛，最后一层释放，用深度损失来微调网络。

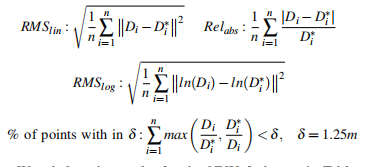
将大模型的最终层移植到初始化的压缩模型。

3、移植模型

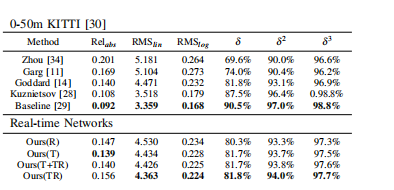


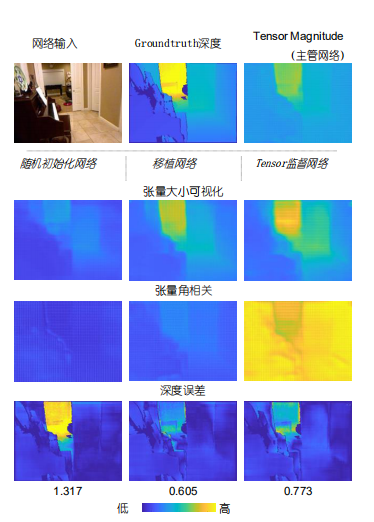
**实验与结果**：





表前三列他们代表error，越小越好。而后三列代表他们内部的百分比越高越好。

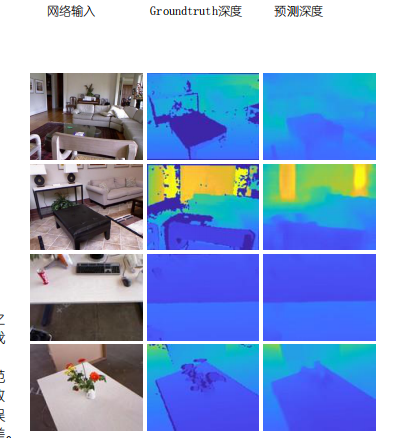
表



观察到随机<Tensor loss <移植

随机模型显然较差的事实凸显了transfer knowledge过程的重要性

从图中可以看出，当使用张量损失训练时，张量角与监督者网络的张量角高度相关，但是激活的幅度相关性较小。



需要注意的一个有趣的事情是，角度与深度误差负相关，即最对齐的张量误差最小，并且在T和TR之间翻转。

**问题**：

它主要提出用于移动平台的压缩移植的模型，文中反复多次强调的transfer knowledge是啥（文中没有什么详细解释）

感受野不怎么理解，增加有什么作用

（我觉得他讲的有点乱）关于他怎么移植这个问题作者没有详细介绍，

关于他的网络架构需要再了解