

# DeepVO: Towards End-to-End Visual Odometry with Deep Recurrent Convolutional Neural Networks

tradition:

feature extraction ---> feature matching ---> motion estimation ---> local optimisation.

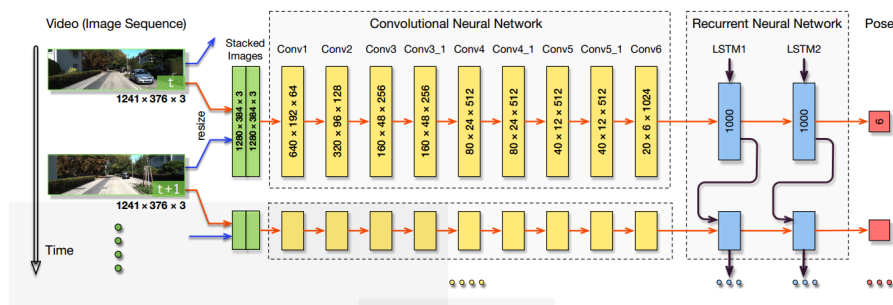
a VO algorithm ideally should model motion dynamics by examining the changes and connections on a sequence of image rather than processing a single image.

we propose a RCNN architecture enabling the DL based VO algorithm to be generalised to totally new environments by using the geometric feature representation learnt by the CNN.

methods based on geometry:

sparse feature based methods and direct methods

It is mainly composed of CNN based feature extraction and RNN based sequential modelling.



CNN提取特征，然后送入RNN，每次训练的sample必须是一个序列，太慢了

具体的做法是：生成一个序列的图像数据：batch × sequence\_length × 6 × w × h

然后resize成 (batch × sequence\_length, 6 × w × h)

cnn提取特征输出大小：(batch × sequence\_length,

feature\_size)

然后resize成 (batch, sequence\_length, feature\_size)

送入rnn 即可，输出为6个参数，x,y,z,roll,pitch,yaw