**1文章信息**

*D-GAN: Deep Generative Adversarial Nets for SpatioTemporal Prediction*

用于时空预测的深度生成对抗网络——D-GAN

Saxena, D., & Cao, J. (2019). D-GAN: Deep Generative Adversarial Nets for Spatio-Temporal Prediction. ArXiv, abs/1907.08556.

**2摘要**

出租车需求、交通流、区域降雨等城市应用的时空(ST)数据具有内在的随机性和不可预测性。最近，提出了基于深度学习的ST预测模型来学习数据的ST特征。然后，这是非常具有挑战性的：

（1）充分学习复杂和非线性的时空关系；

（2）对时空数据的频繁变化建模，时空数据是动态的，随时间变化（不规律），而且被许多因素影响，如变化的天气，交通事故，交通管制，PoI等；

（3）有许多复杂的外部因素能够影响预测的准确性，而且并不能将他们准确的列出来。为了解决以上问题，本文我们提出了一个新颖的基于深度对抗网络的模型（D-GAN），通过无监督学习时空特征表示来更准确地进行时空预测。D-GAN模型采用基础的GAN的结构，学习数据的生成和变化推理。

D-GAN包含两部分：

（1）一个深度时空特征学习网络通过隐形分布建模来对时空关系，语义变化，数据中变化的且不规则的潜在因素建模；

（2）一个聚合模块，用于整合外部因素来更好的推断。据我们所知，以前没有任何工作通过深度隐式生成模型和无监督的方式研究时空预测问题。在两个真实数据集上进行的大量实验表明，该方法比传统的基于深度学习的预测方法获得了更准确的结果。

**3贡献**

本文的贡献有以下三点：

（1）我们提出了一种新颖的基于深度生成对抗网络的模型(命名为D-GAN)，以深度隐式捕捉底层时空数据分布，以实现更准确的时空预测。与现有的时空预测深度模型不同，D-GAN将GAN和VAE结合在一起，在无监督的情况下共同学习时空数据的生成和变分推断。

（2）我们设计了一个通用的聚合模块，对气象数据、时间信息、PoIs数据等异构数据源进行聚合，从而得到更好的预测结果。由于外部因素对时空数据量的变化有显著影响，学习到的潜在表示与时空数据表示合并。D-GAN是高度灵活和可扩展的，因为它可以很容易地扩展到具有多个数据源的新时空数据集。

（3）在两个真实世界基准数据集上的案例研究结果表明，D-GAN比传统的以及基于深度学习的ST预测方法获得了更准确的结果。

**4问题描述**

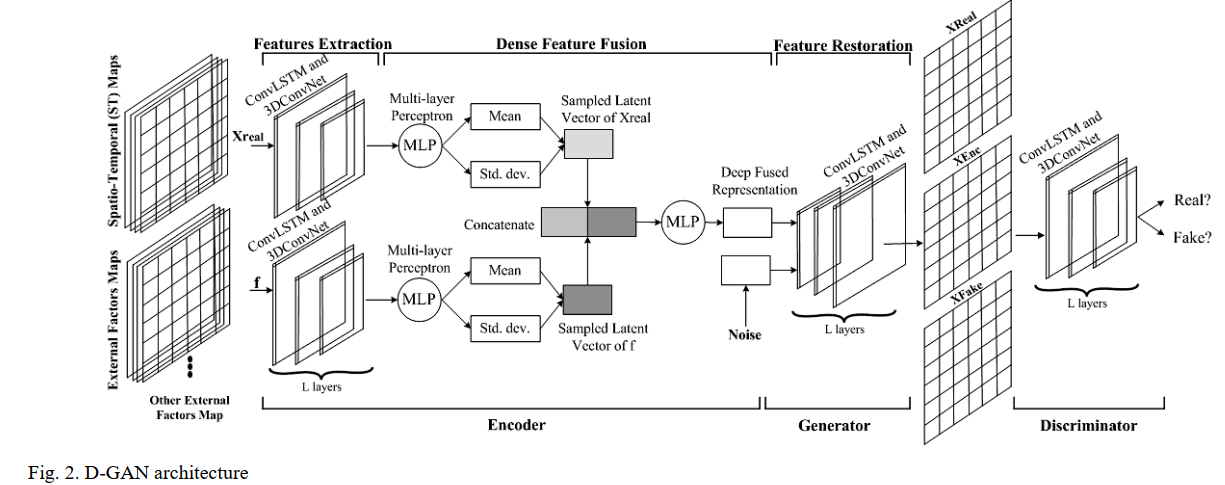
定义1(地区)。对于不同的粒度和语义，有许多定义方式来定义位置。在本研究中，我们将一个城市按照经度和纬度划分为m × n大小的二维网格地图，其中一个网格代表一个区域，一个网格地图称为时空地图。

定义2(测量方法)。对于一个区域，有不同类型的测量方法可以用于不同的时空应用，例如出租车需求和人群流量预测，以及空气质量预测。在本研究中，我们以某市的点播服务预测为例进行研究。因此，我们使用需求值作为测量值。

问题定义：在给出时间间隔t=0,1，……，t-1的历史数据dt的情况下，我们的目的是高精度的预测dt。

**5 D-GAN框架：**

D-GAN的整体网络架构下图所示。



（1）网络体系结构

D-GAN包括四个主要组成部分:

编码器网络(E)、生成器/解码器网络(G)、鉴别器网络(D) 、和外部因素整合(F)。

编码器网络E由一个概率编码器组成，该编码器将数据空间x编码为潜在代码xc。推理网络将参数输出到分布q(xc∣x)。

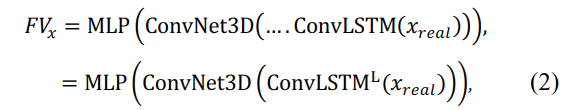
生成器网络由一个概率解码器组成，该解码器学习从表示（或者说潜在代码）xc中重构输入空间x。生成网络将参数输出为似然分布p(xc∣x)。生成器网络通过对抗过程进行训练，生成器（即解码器）学习接近真实数据的分布，而判别器网络则用于区分真实分布的样本和生成器生成的样本。

然后，我们设计了一个通用的融合网络来整合来自不同领域的外部因素与数据。整合模块的数据输入到编码器中，编码器生成外部因素f的潜在代码fc，并将其与提取的特征xc合并。

在D-GAN中，编码器、生成器和判别器使用了ConvLSTM和3D-ConvNet元素的堆叠。其中ConvLSTM神经网络用来捕获时空序列地图中长期的趋势，3D-ConvNet捕获局部空间依赖。ConvLSTM在输入到状态和状态到状态的转换中都有卷积结构。3D卷积能够对跨时间的交通相关性建模。

**编码器和生成器**

在D-GAN中，编码器首先通过ConvLSTM和3DConvNet的多重栈，以及一个多层感知器(MLP)对真实数据进行处理，产生一个精简的特征向量FVx：



其中，XReal是真实数据空间，FVx是提取的xReal的特征向量，L是ConvLSTM的层数。

然后，我们使用具有多变量高斯假设和变分下界损失函数的变分贝叶斯方法，其中模型使用重新参数化技巧明确地计算分布的均值和方差，如下所示：



其中ϵ是辅助独立随机变量，而⊙是元素积。

我们使用KL散度作为正则化项来保证编码器生成的分布与原始数据的分布不会偏离太多。对外部因素f进行同样的处理，将其纳入预测模型。D-GAN支持一个通用的整合网络，将来自不同领域的外部因素与模型整合在一起。我们设计了一个由相似的ConvLSTM层、ConvNet3D层和MLP层组成的特征提取模块。在高斯假设下学习到的辅助特征向量表示为FVfr。

将提取的两个特征向量串联在一块，即FVcat=[FVxrealr,FVfr]，输入到解码器（即生成器）中，也就是说，一个MLP和堆叠的ConvLSTM和3D-ConvNet层按照原来的尺寸重构时空地图。



其中，xEnc是重构的xreal时空地图。此外，噪声向量z传递给生成器作为输入用来生成噪声向量xFake的重构时空地图。

**判别器**

判别器学习判别一个生成的时空地图是来自地面实况还是生成器生成的数据。在D-GAN中，我们将编码FV和其产生的时空地图xEnc联结起来学习潜在编码和数据空间，即yEnc=[xEnc,FVcat]，我们同样生成了yFake=[xFake,FV\*cat]和yreal=[xreal,FVcat]，其中FV\*cat是噪声特征向量。然后，我们使用类似堆叠的ConvLSTM层和ConvNet3D层来实现判别器：



Dout是输入的真实或虚假的预测概率。同时训练生成器和判别器网络，直到判别器不能区分生成器生成的时空图和真实数据生成的时空图。训练后，使用生成器生成与真实数据相似的样本。

（2）损失函数

在本节中，我们将讨论D-GAN的损失函数。我们将其分为两部分:对抗性损失和重建性损失。

**对抗性损失：**对抗性的损失被用来在对抗性训练过程中找到生成器和判别器之间的平衡。在D-GAN中，生成器的对抗性损失为：



我们使用最小二乘损失函数代替基本的生成对抗网络中使用的二元交叉熵来评估差异。

**VAE损失：**包括以下两项损失：（1）KL散度：它测量两个概率分布之间的散度；（2）重建损失(L):它计算地面实况和重建的时空图之间的偏差，以找到网格之间的局部差异。

**6实验**

我们将需求预测表述为一个时空预测问题，其中输入和预测目标都是时空序列。该任务的主要目标是学习一种准确的模型，以预测在每个时间段(时间段可以是一小时、一天或一周)时空地图的每个网格中对特定服务的请求总数。

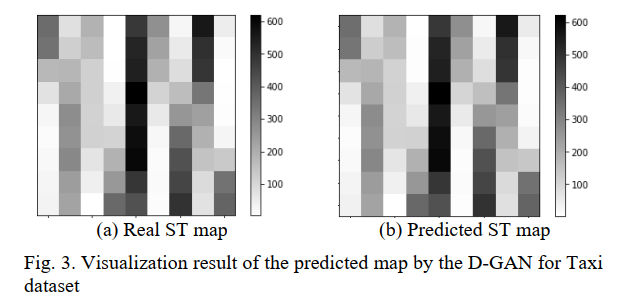
我们使用在纽约市收集的两个大规模数据集：黄色出租车数据集，其中包含从2016年1月1号到2016年6月30号的需求量，以及城市内自行车出行数据集，其中包含从2016年1月1号到2016年1月31号的需求量。

我们将城市表示为一个起点是ls，终点是le的正方形的时空图，正方形区域聚集成9×9的非重叠区域，表示为时空图。需求是一个地区在时间间隔t内出租车/自行车的需求数。一天分为24小时间隔，即时空图代表一小时内的需求。我们将外部因素，如兴趣点(PoI)、天气数据和周末/工作日与历史数据一起用于需求预测。

预处理和参数

我们在训练集上使用最小-最大归一化[0，1]来归一化需求值。在训练之后，我们应用最小-最大变换的逆变换来恢复实际需求值。我们使用四个ConvLSTM层，具有32/16/8/4数量的滤波器和3×3大小的滤波器。我们为ConvLSTM设置了24小时的序列长度。我们在每个ConvLSTM后使用批处理规范化。batchsize为32，epoch为500。我们首先选择90%的数据作为训练数据，剩下的10%用于测试。我们将全连接层的Leaky ReLU作为激活函数。我们使用学习率为0.0001的随机梯度下降算法。此外，为了防止过拟合，我们应用了概率为0.4的dropout方法。我们的模型是用Keras实现的，这是一个深度学习Python库。

结果对比：



上图为真实数据的预处理和D-GAN得到的结果。我们将交通网络中的出租车需求模型化为类似于图像像素矩阵的矩阵，称为时空地图。彩色地图显示了一个网格的出租车请求数量，暗表示请求数量高。

通过与常用基线模型以及与D-GAN的变形的对比，可以显示出本文所构建模型的有效性。

**7创新点**

（1）使用生成对抗网络GAN进行时空数据信息预测，具有新的启发意义，文章设计的D-GAN将有向图形模型(即VAE与GAN)集成为无监督生成模型。D-GAN成功地处理了时空相关和存在于时空数据中的随机元素。

（2）设计了整合模块，融合外部因素信息，提高了D-GAN的估计能力。