

基于改进生成式对抗网络的电气数据升频重建方法

李富盛, 林丹, 余涛, 王克英, 吴毓峰, 杨家俊

(华南理工大学电力学院, 广东省广州市 510640)

摘要: 高频电气数据是提高电网态势感知准确度、监测水平和辅助服务质量等的数据库之一,但是,传统重建算法难以实现高精度的数据重建。因此,文中利用改进生成式对抗网络将低频电气数据重建为高频。通过将时序数据转化为电气图像,实现神经网络方法对电气图像特征的高效提取。利用基于深层残差网络的生成器和改进的残差块结构,提高生成器的特征学习能力。此外,生成器损失函数考虑真实样本与生成样本在低维或高维特征的差别。以公开数据集为例进行算法验证,验证结果表明,相比于传统重建方法,所提方法具有更高的峰值信噪比、结构相似性和更低的平均绝对误差、平均绝对误差百分数,以及更高的高频细节还原度、重建精度,能够对不同数据集实现泛化。

关键词: 升频重建; 低频电气数据; 生成式对抗网络; 深度残差网络; 图像特征; 数据驱动

0 引言

物联网的目的之一是实现不同对象间的物理连接与信息交互^[1],对数据采集、数据清洗^[2]、数据通信^[3]、数据高级应用^[4]等方面提出更高的要求。信息与通信技术是提高电网自动化水平的重要手段,也使电网产生大量高频和低频数据^[5]。相比于低频电气数据,高频电气数据具有更丰富的多源异构的特征,能够更清晰地反映用户用电行为,形成更准确的用户画像和标签体系。其次,高频电气数据涉及更多的维度,能够提高相关性挖掘能力,提升预测水平。另一方面,由于通信和储存技术限制,高频数据常通过压缩感知技术转化为低频数据,应用前再进行升频重建(frequency-increased reconstruction, FIR),还原数据细节^[6]。同时,如果数据采集与监控(SCADA)系统^[7]将高频数据压缩为低频后再传输,则能在受到信息攻击后通过数据重建而降低数据损坏程度。综上,升频重建是提高数据质量和挖掘电网数据潜在价值的支撑技术之一。

电气领域中已有的重建算法着力于缺失片段重建问题^[8-10],而不是提高数据的采样频率(即升频重建)。例如,文献[8]利用一维电气数据和基于浅层

神经网络的生成式对抗网络(GAN)^[11]实现高精度的缺失片段重建,但未讨论该模型在升频重建问题的适用性。缺失片段重建问题一般是整体完整、局部缺失,为了保证重建精度,缺失片段的数据量一般不超过原始数据总量的50%,即重建倍数不超过2倍^[8]。而升频重建的目的在于提高整个数据集的采样频率,重建倍数一般不低于2倍。

目前,电气领域常用插值方法对一维时序电气数据进行升频重建^[12-13],但重建精度不高,且缺乏高频细节。图像领域的超分辨率重建方法主要有3类:基于插值的方法^[14]、基于建模的方法^[15-16]和基于学习^[17]的方法,但这些方法的高频细节还原度不高。近年来,GAN在图像、文本、能源等领域取得突破性进展^[18-21]。因此,有必要设计一种将一维时序电气数据转化为电气图像的方法,利用GAN在图像超分辨率重建问题的优势,提高电气数据的升频重建精度。同时,电气数据升频重建涉及数据从低维向高维转化,是一个具有大量可能解的高病态逆问题^[22-23],有必要设计适用于该问题的GAN的生成器、判别器和损失函数。

综上,为了解决低频电气数据的升频重建问题,提高重建精度和高频细节还原度,本文设计了电气数据转化为电气图像的方法,引入了GAN,损失函数考虑对抗损失、实际损失、感知损失和电气图像的低维特征或高维特征,生成器引入了深层残差网络并对残差块(residual block)进行改进。最后,以国外公开可用的数据集为例,验证所提方法的有效性。

收稿日期: 2021-02-11; 修回日期: 2021-03-02。

上网日期: 2021-08-25。

国家自然科学基金资助项目(U2066212);广东省普通高校基础研究与应用基础研究重点项目(2018KZDXM001)。

1 基于GAN的电气数据升频重建

1.1 新型数据通信机制

近年来智能电表项目被大力推广,但国际公开的电力数据集中大部分的采集频率较低,采集周期级别大多为15 min、30 min、1 h等^[24],只有少数达到了1 min或者1 s级别^[25-28]。由于电网设备数量大、类型多,导致电气测量数据规模大,高频采样面临投入成本高、通信难度大、数据存储成本高等限制,所以电网已有的数据集一般是低频。传统的数据通信机制是直接传输低频或高频数据。直接传输低频数据会忽略部分数据细节,传输高频数据会增大通信和储存的压力。

显然,该模式难以满足数据量激增的现代电网的要求。因此,本文提出一种新型数据通信机制,高频数据在数据发送端压缩为低频,低频数据在数据接收端就地进行升频重建。该模式解决了传统模式的不足:一是分布式的数据处理方式减轻了数据中心的计算和存储压力;二是通信网只需传输低频数据,相同带宽下可以传输更多的数据集,通信效率高。

1.2 数据升频重建的总体框架

本文设计了低频电气数据的升频重建框架,如图1所示。

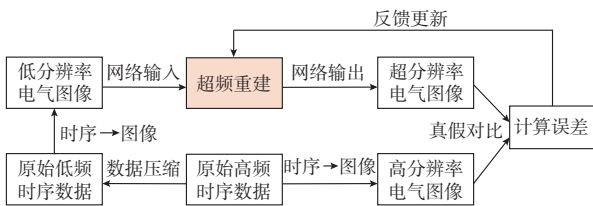


图1 数据升频重建的总体框架
Fig. 1 Overall framework of data FIR

首先对原始高频数据集下采样得到低频数据集,然后将时序形式的电气数据转化为电气图像,利用电气图像进行升频重建,最后比较升频数据集与原始高频数据集的误差,根据误差反馈更新重建模型的参数。本文中“高频”不是绝对的,而是相对的概念。升频重建是指将数据采样频率以一定的重建倍数进行提高。本文定义仿真用的原始数据为高频数据,对原始数据进行下采样得到低频数据,重建后的数据称为升频数据,升频数据与高频数据的采样率相同。此外,原始数据采样率的高低不是限制本文方法的因素。

1.3 电气数据转化为电气图像

为了充分利用先进机器学习方法在图像特征提取上的优势,本文设计了电气数据转化为电气图像的通用方法。设高频电气数据集有 L 个采样时刻,

每个采样时刻有 n 种类型的电气数据,包括有功功率、无功功率、电压、电流、频率、功率因数等,则高频电气数据集可以表示为 $L \times n$ 的矩阵。每张电气图像的 n 个通道分别对应电气数据的 n 种类型。图2和附录A表A1分别为时序数据转化为电气图像的示意图和步骤。图2中, l 为电气图像每个通道的边长。电气图像形式不改变时序形式的数据大小,仅改变数据的排列位置。电气图像的颜色深浅与时序数据的大小正相关,其纹理变化与时序数据的变化趋势正相关,根据电气图像能够从视觉上比较高频细节还原度。

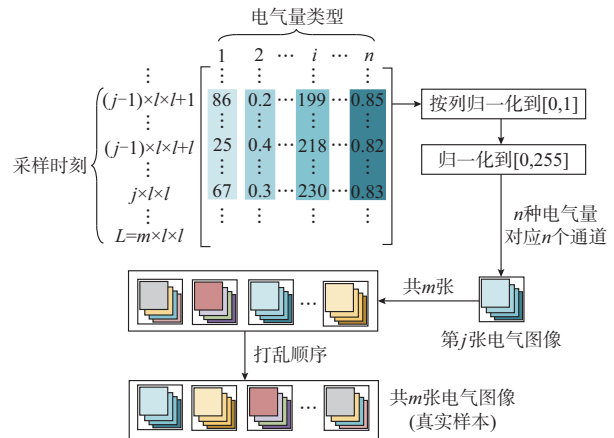


图2 时序数据转化为电气图像示意图
Fig. 2 Schematic diagram of transforming time series data into electrical image

1.4 GAN

GAN通过生成器生成超分辨率电气图像,再通过判别器从大量可能解中筛选出与实际情况差异最小的生成样本,从而解决高病态逆问题,实现高精度重建^[29]。GAN只是一种框架,其生成器和判别器的内部算法、连接方式、损失函数等在不同的问题上会有较大的区别。

设已有高分辨率电气图像集、低分辨率电气图像集和超分辨率电气图像集分别为 x^{HR} 、 x^{LR} 和 x^{SR} ,其数据单元分别为 x_j^{HR} 、 x_j^{LR} 和 x_j^{SR} ,其中 x_j^{SR} 表示生成器 G_{θ_G} 以 x^{LR} 为输入得到的生成样本, x_j^{SR} 也可以表示为 $G_{\theta_G}(x^{LR})$,其中 θ_G 为生成器参数。判别器输入为生成样本或真实样本,输出为输入被判定为真实样本的概率,为 $D_{\theta_D}(G_{\theta_G}(x^{LR}))$ 或 $D_{\theta_D}(x^{HR})$,其中 θ_D 为判别器参数, $D_{\theta_D}(x^{SR})$ 或 $D_{\theta_D}(x^{HR})$ 分别表示判别器 D_{θ_D} 以 x^{SR} 或 x^{HR} 为输入得到的输出。基于GAN的升频重建架构如附录A图A1所示。

1.5 生成器

1.5.1 生成器结构

电力系统是巨维且复杂的,其电气数据具有高

维潜在特征。一般思路是引入神经网络对电气数据进行特征挖掘,但是由于神经网络存在梯度消失问题,本文引入学习能力更强的深度残差网络(DRN)作为生成器。DRN内部的残差块使用了残差跳跃式的结构,打破了神经网络某一层的输出只能给下一层作为输入的惯例,使某一层的输出可以跨过几层作为后面某一层的输入,缓解了神经网络深度增加带来的梯度消失问题,并能通过叠加多层残差块提高模型的特征学习能力和准确率^[30]。生成器结构如图3所示。其中, K 为重建倍数。为了加速训练,利用VGG-19模型^[31]的预训练结果初始化DRN。

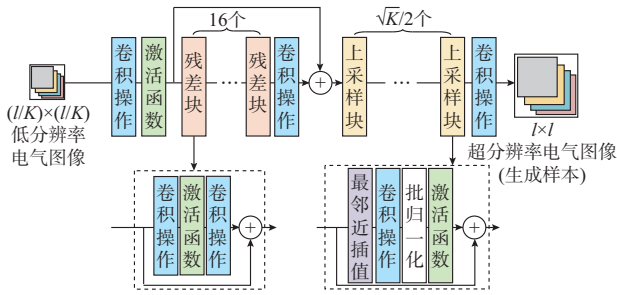


图3 数据升频重建的生成器结构

Fig. 3 Structure of generator for data FIR

在一定网络层数下,由于残差结构的引入仅增加加法运算,深度置信网络(DBN)与深度卷积网络(DCN)^[32]具有相同的算法复杂度^[30]。由于残差块的大部分参数集中在批归一化层,因此残差块去掉批归一化层有助于降低计算复杂度。原始残差块和改进残差块结构如附录A图A2所示。设深层残差网络具有 R 个残差块,每个原始残差块具有2个批归一化层,每个批归一化层的输入特征图的个数、通道数、长、宽分别为 m 、 n 、 l_r 、 l_r 。每次迭代需要计算每个残差块的每个特征图的每个通道的均值和标准差各 $2Rmn$ 次,即批归一化层需要优化的参数数量为 $4Rmn$ 个。对每个像素进行归一化处理,共需要归一化计算 $\sum_{r=1}^R 4mnl_r l_r$ 次。相比于基于原始残差块的深度残差网络,残差块去掉批归一化层能够节省约40%的内存^[33]。因此,所提生成器模型的残差块去掉批归一化层,仅保留卷积操作、激活函数和残差跳跃式结构。

1.5.2 生成器损失函数

为了提高重建精度和高频细节还原能力,生成器损失函数设计为以下几种损失函数的加权组合。

为了使得生成器能够欺骗判别器,生成对抗损失 l_{Gen} 考虑了判别结果,计算如下:

$$l_{Gen} = - \sum_{j=1}^m D_{\theta_D}(G_{\theta_G}(x_j^{LR})) \quad (1)$$

实际损失为生成样本和真实样本的均方误差(MSE),定义为 l_{MSE} ,其表达式为:

$$l_{MSE} = \frac{1}{l^2} \sum_{j=1}^m \sum_{w=1}^l \sum_{h=1}^l (x_{j,w,h}^{HR} - x_{j,w,h}^{SR})^2 \quad (2)$$

式中: w 和 h 分别为每一个通道中矩阵的宽和高; $x_{j,w,h}^{HR}$ 和 $x_{j,w,h}^{SR}$ 分别为第 j 个真实样本、生成样本的每个通道矩阵中宽为 w 和高为 h 的像素点的数值。

使用实际损失训练生成器会导致生成样本过于平滑,重建结果过于保守,不利于高频细节的还原。因此,本文引入感知损失衡量生成样本和真实样本在局部特征上的差异。具体地,考虑到VGG-19模型具有提取电气图像局部特征的能力,本文在已经训练好的VGG-19模型中分别输入生成样本和真实样本,并提取电气图像的特征图 $\varphi_{a,b}$,其中 $\varphi_{a,b}$ 的含义为VGG-19模型在第 a 次池化操作之前得到的第 b 个特征图。本文将真实样本特征图和生成样本特征图在对应位置上各像素点数值的均方误差定义为感知损失 l_{VGG} 。对于基于低维特征的本文方法,感知损失使用特征图 $\varphi_{2,1}$ 进行计算,对于基于高维特征的本文方法,感知损失使用特征图 $\varphi_{5,4}$ 进行计算。 l_{VGG} 的计算公式为:

$$l_{VGG} = \frac{1}{l^2} \sum_{j=1}^m \sum_{w=1}^l \sum_{h=1}^l (\varphi_{a,b}(x_{j,w,h}^{HR}) - \varphi_{a,b}(x_{j,w,h}^{SR}))^2 \quad (3)$$

式中: $\varphi_{a,b}(x_{j,w,h}^{HR})$ 和 $\varphi_{a,b}(x_{j,w,h}^{SR})$ 分别为第 j 个真实样本特征图和生成样本特征图的每个通道矩阵中宽为 w 、高为 h 的像素点的数值。

综上,生成器损失函数为:

$$l_G = \lambda_1 l_{Gen} + \lambda_2 l_{MSE} + \lambda_3 l_{VGG} \quad (4)$$

式中: λ_1 、 λ_2 、 λ_3 为加权系数,为经验值,取 $\lambda_1 = 10^{-3}$, $\lambda_2 = 1$, $\lambda_3 = 2 \times 10^{-4}$ 。

1.6 判别器

1.6.1 判别器结构

DCN作为判别器,能够实现真实样本与生成样本的有效分类。生成器实现以假乱真,而判别器实现去伪存真。通过二者的对抗博弈,生成器的生成能力和判别器的判别能力会同步提高,使得生成样本更接近真实样本,最终实现更高的重建精度。判别器结构如附录A图A3所示。

1.6.2 判别器损失函数

为了使得判别器能够辨别真假,判别对抗损失 l_D 考虑了判别结果,计算式如下:

$$l_D = \sum_{j=1}^m D_{\theta_D}(G_{\theta_G}(x_j^{LR})) - \sum_{j=1}^m D_{\theta_D}(x_j^{HR}) \quad (5)$$

2 仿真实验

本章通过3个案例验证所提升频重建方法的效果。案例1对比本文方法与其他方法的高频细节还

原度;案例2对比本文方法与其他方法的重建精度。案例3研究本文方法对同一个地理位置的不同数据集的泛化性。

迭代次数设为5000次,使用Adam优化器,初始学习率设为 1×10^{-4} ,Adam优化器的一阶矩估计的指数衰减率设为0.9,每500次迭代后学习率变为原来的1/10。

2.1 数据准备

2.1.1 缺失值填补

截取数据量比较完整的片段作为实验数据。由于每个缺口的缺失数量对数据总量而言可以忽略,本文使用线性插值方法进行缺失值填补。

2.1.2 归一化处理

为了提高电力图像的可视化效果,先利用最大最小归一化方法将电气测量数据的取值范围调整为 $[0,1]$,然后所有数据乘以255,变换到图像处理的常用取值范围 $[0,255]$ 。相同类型电气数据在归一化和反归一化中使用相同的最大值和最小值。

2.1.3 仿真用数据集

以国外公开可用的数据集I-BLEND^[28]为例,对3个案例进行仿真验证。I-BLEND是某大学校园内不同建筑的用电数据集,建筑类型包括教学楼、图书馆、设备房、报告厅、男生宿舍、女生宿舍和餐厅,采样时长为52个月(每座建筑的单个电气量的数据集为200万数量级),电气参数包括电压、电流、功率、频率和功率因数,采样频率为1min。仿真用的数据集片段截取为当地时间2014-01-08T00:00到2015-05-02T23:59,共有691200个采样点,建筑类型为教学楼、男生宿舍、女生宿舍和餐厅,电气参数为有功功率、电流和电压。

本文取低频和高频电气测量数据的采样频率分别为15min/次和1min/次。为了方便构造电气图像,在高频电气测量数据的每15个数据后添加一个0值,使重建倍数 K 变为16, \sqrt{K} 的结果为整数。加零操作不会对低分辨率电气图像造成影响。根据上述处理,每64h的电气测量数据转化1张电气图像。本文仿真中,高分辨率电气图像和重建后的超分辨率电气图像的大小均为 $64 \times 64 \times 3$,低分辨率电气图像的大小为 $16 \times 16 \times 3$ 。每个数据集都转化为180张高分辨率电气图像和180张低分辨率电气图像,并按照8:1的比例划分训练集和验证集。

2.2 数据重建精度评估指标

2.2.1 时序数据层面

平均绝对误差(MAE)的范围为 $[0, +\infty)$,MAE越小,误差越小,计算式如下:

$$I_{MAE} = \frac{1}{T} \sum_{t=1}^T |y_t^* - y_t| \quad (6)$$

式中: I_{MAE} 为MAE值; y_t^* 、 y_t 分别为验证集的原始时序电气数据和重建时序电气数据,其中 t 表示验证集的第 t 个数据; T 为验证集的数据总量。

平均绝对误差百分数(MAPE)的范围为 $[0, +\infty)$,MAPE越小,误差越小,计算式如下:

$$I_{MAPE} = \frac{100\%}{T} \sum_{t=1}^T \left| \frac{y_t^* - y_t}{y_t} \right| \quad (7)$$

式中: I_{MAPE} 为MAPE值。

2.2.2 图像层面

峰值信噪比(PSNR)^[34]是重建领域的常见评价指标。PSNR常用MSE定义,即基于像素点的误差进行图像质量评价。PSNR通过非线性运算将MSE转化为单位为dB,范围 $(-\infty, +\infty)$ 的值。根据文献[23],当前技术水平下可认为PSNR大于20dB为较高的重建精度。PSNR的计算式为:

$$I_{PSNR} = 10 \lg \frac{255^2}{I_{MSE}} \quad (8)$$

$$I_{MSE} = \frac{1}{l^2} \sum_{w=1}^l \sum_{h=1}^l (x_{j,w,h}^{HR} - x_{j,w,h}^{SR})^2 \quad (9)$$

式中: I_{PSNR} 为PSNR值; I_{MSE} 为MSE值。

电气图像是高度结构化的,相邻像素之间有很强的关联性。结构相似性(SSIM)^[35]可以抵消MAE和MAPE无法衡量图像结构相似性的缺陷。SSIM从图像组成的角度将图像信息定义为亮度、对比度和结构3个不同因素的组合,用均值作为亮度的估计,标准差作为对比度的估计,协方差作为结构相似程度的评估指标,详细推导参考文献[35]。SSIM的简化计算式如下:

$$I_{SSIM} = \frac{2\mu_{x_j^{HR}}\mu_{x_j^{SR}} + C_1}{\mu_{x_j^{HR}}^2 + \mu_{x_j^{SR}}^2 + C_1} \frac{2\sigma_{x_j^{HR}x_j^{SR}} + C_2}{\sigma_{x_j^{HR}}^2 + \sigma_{x_j^{SR}}^2 + C_2} \quad (10)$$

式中: I_{SSIM} 为SSIM值; $\mu_{x_j^{HR}}$ 和 $\mu_{x_j^{SR}}$ 分别为 x_j^{HR} 和 x_j^{SR} 的平均值; $\sigma_{x_j^{HR}}$ 和 $\sigma_{x_j^{SR}}$ 分别为 x_j^{HR} 和 x_j^{SR} 的标准差; $\sigma_{x_j^{HR}x_j^{SR}}$ 为 x_j^{HR} 和 x_j^{SR} 的协方差; C_1 和 C_2 为常数,本文取 $C_1 = (0.01 \times 255)^2$, $C_2 = (0.03 \times 255)^2$ 。

由1.3节可知,电气时序数据的变化趋势和数值分别对应电气图像的纹理和色调。SSIM值的范围是 $[0,1]$,越大表示电气图像的纹理与色调越接近,即时序形式的电气数据的变化趋势与数值越接近。

2.3 高频细节还原度评估

对双三次插值(BI)^[36]、传统GAN^[8]和深度卷积神经网络(DCNN)^[31]以及本文基于低维、高维特征的数据重建方法进行高频细节还原度评估。

2.3.1 电气图像

以教学楼数据集为例,附录A图A4为电气图

像形式的重建结果。高分辨率电气图像在视觉上具有丰富的高频细节和清晰的纹理。插值重建的电气图像虽然像素增大了,但是在视觉上趋于平滑,重建效果相对保守,表明了插值重建方法不具备高频细节还原能力。采用传统GAN和DCNN方法重建的电气图像比插值方法有更清晰的纹理,但是与原始高分辨率电气图像相比,整体色调以及局部纹理细节仍有明显差别。基于低维特征或高维特征的超分辨率电气图像的高频细节还原度都很高,其生成的纹理和色调与高分辨率电气图像高度相似,但是基于高维特征的电气图像比基于低维特征的电气图像具有更丰富的纹理。

2.3.2 时序数据变化曲线

以教学楼数据集为例,将验证集的电气图像还原为时序形式的电气测量数据,数据重建结果如图4所示。采用本文方法得到的重建曲线与真实曲线相比,除了个别区间的峰值较低或谷值较高外,2条曲线的变化趋势基本相同,整体重建效果良好。

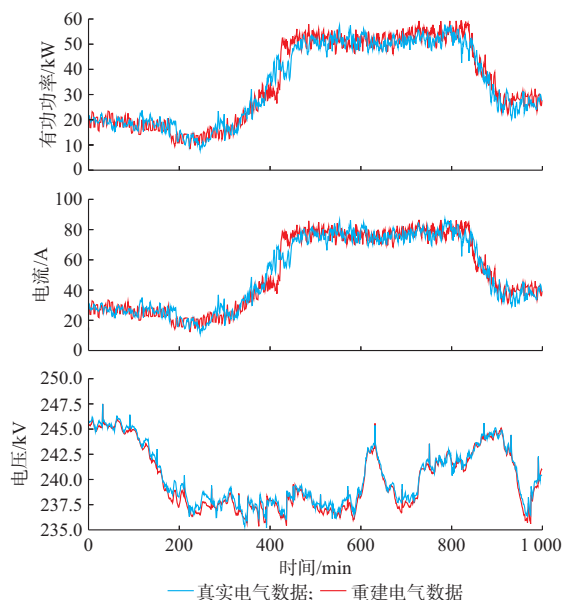


图4 数据FIR结果
Fig. 4 Results of data FIR

2.4 数据重建精度评估

由于高频细节的多样性和随机性,侧重于对点与点之间的片段进行重建会导致无穷多解。因此实现重建应该分两步走,先重建整体相似分布,再还原高频细节。GAN能够学习原始数据的潜在分布,有助于对解空间进行降维,再加上本文设计的提高高频细节还原能力的手段,最终实现高精度重建。以下从分布相似性和评估指标方面进行重建精度评估。

2.4.1 分布相似性

本研究验证了重建数据具有与真实数据相同的统计特性。图5为关于教学楼有功功率的真实数据和重建数据的累积分布函数(CDF),即小于等于当前数据值的所有数据的概率分布。对于给定的数据集,累积分布函数是唯一的,不会丢失任何数据信息。由图5可见,重建数据的累积分布除了峰值较低或谷值较高外,其他部分与真实数据的累积分布基本相同。此外,本研究利用JS距离^[37]计算真实数据与重建数据的累积分布的相似程度。JS距离是信息论中常用于衡量概率分布相似性的指标,取值范围为 $[0,1]$,越接近0表示两分布差异越小。本实验的JS距离计算结果为0.18,接近于0,表明真实数据和重建数据的累积分布有小部分存在差别,但整体相似度较高。

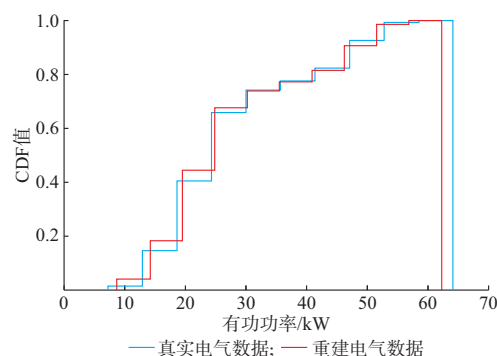


图5 CDF评估结果
Fig. 5 CDF assessment results

2.4.2 评估指标

根据式(7)至式(11),计算验证集的评估指标。表1为在单次验证中各项指标的平均评估值,表2为在单次验证中各图像的各个指标评估值。可见,针对不同建筑的电气数据集,插值重建方法的PSNR和SSIM的数值都比较小,表明该方法无法有效挖掘电气数据的数据结构关系,高频细节还原能力不足。插值重建方法的MAE和MAPE的数值都比较大,表明该方法重建精度较低,无法准确还原高随机性的高频细节。传统GAN方法的PSNR不是特别突出,SSIM比较高,表明传统GAN方法能够在一定程度上挖掘电气数据的数据结构关系,但是没有经过改进的GAN的高频细节还原能力不理想。传统GAN方法和DCNN的MAE和MAPE在3种方法中处于中间水平,表明该方法具有一定的重建精度。相对地,本文方法的PSNR比较高,表明该方法的重建结果失真比较小。超分辨率重建方法的SSIM接近1,表明该方法能够有效地重建电力图像的纹理和色调,即能够学习到电气测量数据的结构关系,进而有效地还原高频细节。本文方法的

MAE和MAPE与其他2种方法相比最小,重建精度最高。对于本文方法,基于高维特征的重建结果比基于低维特征的重建结果具有更高的PSNR、SSIM和更低的MAE、MAPE,表明基于高维特征的数据重建方法更适用于本研究的数据集。

表1 不同方法的数据FIR精度
Table 1 Data FIR accuracy of different methods

重建方法	I_{PSNR}/dB	I_{SSIM}	I_{MAE}/kW	$I_{MAPE}/\%$
BI	13.72	0.48	8.92	15.85
传统GAN	18.21	0.91	6.75	12.17
DCNN	17.85	0.79	4.58	9.22
本文方法(低维特征)	25.36	0.96	3.15	5.87
本文方法(高维特征)	26.14	0.97	2.46	4.32

表2 基于高维特征的数据FIR精度
Table 2 Data FIR accuracy based on high-dimensional features

样本序号	I_{PSNR}/dB	I_{SSIM}	I_{MAE}/kW	$I_{MAPE}/\%$
1	26.23	0.98	2.32	4.09
2	26.01	0.93	2.43	4.32
3	26.17	0.98	2.31	4.01
4	26.23	0.98	2.53	4.24
5	25.41	0.96	3.05	5.42
6	25.77	0.97	2.42	4.35
7	25.26	0.97	3.55	6.58
8	25.62	0.97	2.28	3.89
9	25.80	0.97	2.42	3.93
10	25.95	0.97	2.38	4.19
11	26.77	0.98	2.12	3.72
12	26.11	0.98	2.37	4.18
13	27.09	0.97	2.26	3.95
14	25.58	0.97	3.42	6.22
15	26.13	0.96	2.39	4.17
16	26.35	0.97	2.25	4.05
17	25.99	0.97	2.27	3.96
18	26.07	0.95	2.32	4.10
19	26.77	0.99	2.12	3.56
20	27.49	0.99	2.04	3.48

2.5 重建模型在区域内不同数据集间的泛化性

为了研究相同地理位置下不同建筑的电气数据单独训练得到的模型对于其他建筑的泛化性,本节将上文中教学楼、男生宿舍、女生宿舍和餐厅的生成器作为已训练的重建模型,将4个建筑的验证集交叉输入到4个重建模型中,得到新的超分辨率电气图像。附录A图A5为某一验证集在区域内不同数据重建模型下的升频重建结果。结果表明,教学楼和女生宿舍的验证集,在教学楼重建模型和女生宿

舍重建模型上都能恢复相似的纹理和色调。类似地,男生宿舍和餐厅的验证集,在男生宿舍重建模型和餐厅重建模型上都能够恢复相似的纹理和色调。不同建筑模型之间存在泛化性是因为GAN从不同建筑数据集学到了相似的潜在分布。用电数据的潜在分布与用电行为是强关联的,因此,I-BLEND数据集中,教学楼和女生宿舍具有相似的用电行为,男生宿舍和餐厅具有相似的用电行为。

除了视觉上对高频细节还原度的比较,附录A表A2展示了区域内不同数据集的重建精度评估结果。教学楼和女生宿舍验证集,在教学楼重建模型和女生宿舍重建模型上均具有较高的PSNR、SSIM和较低的MAE、MAPE。男生宿舍和餐厅验证集,在男生宿舍重建模型和餐厅重建模型上,均具有较高的PSNR、SSIM和较低的MAE、MAPE。结果表明,本文方法能够在相同区域内的不同数据集间进行泛化。

3 结语

本文将改进GAN应用于低频电气数据的升频重建问题,其主要贡献如下:

1)设计了时序形式的电气数据转化为电气图像的方法,利用先进神经网络方法实现了对电气图像特征的高效挖掘;

2)通过采用基于深层残差网络的生成器解决了深度神经网络的梯度消失问题,通过设计残差块的结构降低了计算复杂度,并在生成器损失函数中考虑电气图像的低维特征或高维特征,提高了生成器学习能力;

3)本文方法是基于数据驱动的无监督训练方法,具有一定的通用性,仿真结果表明该方法在相同的模型参数下能够对不同的数据集实现高精度升频重建,且训练好的重建模型对同一个地理位置的不同数据集具有可泛化性。

本文未将数据压缩与重建方法进行综合考虑,且未研究对不同地区的电气数据集均具有泛化性的升频重建方法,未来需要进一步加深该方面研究。

附录见本刊网络版(<http://www.aeps-info.com/aeps/ch/index.aspx>),扫英文摘要后二维码可以阅读网络全文。

参考文献

- [1] 程乐峰,余涛,张孝顺,等.机器学习在能源与电力系统领域的应用和展望[J].电力系统自动化,2019,43(1):15-43.
CHENG Lefeng, YU Tao, ZHANG Xiaoshun, et al. Machine learning for energy and electric power systems: state of the art and prospects[J]. Automation of Electric Power Systems, 2019, 43(1): 15-43.

- [2] 代杰杰,宋辉,杨祎,等.基于栈式降噪自编码器的输变电设备状态数据清洗方法[J].电力系统自动化,2017,41(12):224-230. DAI Jiejie, SONG Hui, YANG Yi, et al. Cleaning method for status data of power transmission and transformation equipment based on stacked denoising autoencoders [J]. Automation of Electric Power Systems, 2017, 41(12): 224-230.
- [3] GUNGOR V C, SAHIN D, KOCAK T, et al. Smart grid technologies: communication technologies and standards [J]. IEEE Transactions on Industrial Informatics, 2011, 7(4): 529-539.
- [4] 余涛,程乐峰,张孝顺.基于信息-物理-社会系统融合和群体机器学习的弱中心化微元网:理论研究与关键科学问题分析[J].中国科学(技术科学),2019,49(12):1541-1569. YU Tao, CHENG Lefeng, ZHANG Xiaoshun. The weakly-centralized web-of-cells based on cyber-physical-social systems integration and group machine learning: theoretical investigations and key scientific issues analysis [J]. Scientia Sinica (Technologica), 2019, 49(12): 1541-1569.
- [5] 李彬,卢超,景栋盛,等.负载与风险联合均衡的电力通信路由优化算法[J].中国电机工程学报,2019,39(9):2713-2722. LI Bin, LU Chao, JING Dongsheng, et al. An optimized routing algorithm with load and risk joint balance in electric communication network[J]. Proceedings of the CSEE, 2019, 39(9): 2713-2722.
- [6] LAYA D, DINESH G, BABJI S. Neural compression: a machine learning approach to compress high frequency measurements in smart grid [J]. Applied Energy, 2020, 257: 113966.
- [7] 倪明,颜洁,柏瑞,等.电力系统防恶意信息攻击的思考[J].电力系统自动化,2016,40(5):148-151. NI Ming, YAN Jie, BO Rui, et al. Power system cyber attack and its defense [J]. Automation of Electric Power Systems, 2016, 40(5): 148-151.
- [8] 王守相,陈海文,潘志新,等.采用改进生成式对抗网络的电力系统量测缺失数据重建方法[J].中国电机工程学报,2019,39(1):56-64. WANG Shouxiang, CHEN Haiwen, PAN Zhixin, et al. A reconstruction method for missing data in power system measurement using an improved generative adversarial network [J]. Proceedings of the CSEE, 2019, 39(1): 56-64.
- [9] 杨挺,李扬,何周泽,等.基于矩阵填充的泛在电力物联网电能质量数据修复算法[J].电力系统自动化,2020,44(2):13-22. YANG Ting, LI Yang, HE Zhouze, et al. Matrix completion theory based recovery algorithm for power quality data in ubiquitous power Internet of Things [J]. Automation of Electric Power Systems, 2020, 44(2): 13-22.
- [10] 梁寿愚,方文崇,王瑾,等.基于图规则化低秩矩阵恢复的用电数据修复与异常检测[J].电力系统自动化,2019,43(21):221-228. LIANG Shouyu, FANG Wenchong, WANG Jin, et al. Refinement and anomaly detection for power consumption data based on recovery of graph regularized low-rank matrix [J]. Automation of Electric Power Systems, 2019, 43(21): 221-228.
- [11] GOODFELLOW I J, POUGET-ABADIE J, MIRZA M, et al. Generative adversarial networks [J]. Advances in Neural Information Processing Systems, 2014(3): 2672-2680.
- [12] ZHANG F, CHENG L, LI X, et al. Application of a real-time data compression and adapted protocol technique for WAMS [J]. IEEE Transactions on Power Systems, 2015, 30(2): 653-662.
- [13] MIRANDA V, KRSTULOVIC J, KEKO H, et al. Reconstructing missing data in state estimation with autoencoders [J]. IEEE Transactions on Power Systems, 2012, 27(2): 604-611.
- [14] WANG L F, XIANG S M, MENG G F, et al. Edge-directed single-image super-resolution via adaptive gradient magnitude self-interpolation [J]. IEEE Transactions on Circuits & Systems for Video Technology, 2013, 23(8): 1289-1299.
- [15] BABACAN S D, MOLINA R, KATSAGGELOS A K. Variational Bayesian super resolution [J]. IEEE Transactions on Image Processing, 2011, 20(4): 984-999.
- [16] LUKES T, KRÍŽEK P, SVINDRYCH Z, et al. Three-dimensional super-resolution structured illumination microscopy with maximum a posteriori probability image estimation [J]. Optics Express, 2014, 22(24): 29805.
- [17] LU X, YUAN Y, YAN P. Alternatively constrained dictionary learning for image super resolution [J]. IEEE Transactions on Cybernetics, 2014, 44(3): 366-377.
- [18] KONSTANTINOS B, NATHAN S, DAVID D, et al. Unsupervised pixel-level domain adaptation with generative adversarial networks [EB/OL]. [2017-08-23]. <https://arxiv.org/abs/1612.05424v1>.
- [19] HAN Z, TAO X, HONGSHENG L. StackGAN: text to photo-realistic image synthesis with stacked generative adversarial networks [C]// Proceedings of 2017 IEEE International Conference on Computer Vision (ICCV), October 22-29, 2017, Venice, Italy: 5908-5916.
- [20] 孙秋野,刘月,胡旌伟,等.基于GAN的非侵入式自能源建模[J].中国电机工程学报,2020,40(21):6784-6794. SUN Qiuye, LIU Yue, HU Jingwei, et al. Non-intrusive we-energy modeling based on GAN technology [J]. Proceedings of the CSEE, 2020, 40(21): 6784-6794.
- [21] 董晓翀,孙英云,蒲天骄.基于条件生成对抗网络的可再生能源日前场景生成方法[J].中国电机工程学报,2020,40(17):5527-5536. DONG Xiaochong, SUN Yingyun, PU Tianjiao. Day-ahead scenario generation of renewable energy based on conditional GAN [J]. Proceedings of the CSEE, 2020, 40(17): 5527-5536.
- [22] LUCAS A, LOPEZ-TAPIA S, MOLINA R, et al. Generative adversarial networks and perceptual losses for video super-resolution [J]. IEEE Transactions on Image Processing, 2019, 28(7): 3312-3327.
- [23] LEDIG C, THEIS L, HUSZAR F, et al. Photo-realistic single image super-resolution using a generative adversarial network [C]// IEEE Conference on Computer Vision & Pattern Recognition, July 21-26, 2017, Honolulu, USA: 1063-6919.
- [24] Smart metering equipment technical specifications: version 2 [EB/OL]. [2020-10-21]. <http://bit.ly/1401o61>.
- [25] UTTAMA NAMBI A S N, REYES LUA A, PRASAD V R. LocED: location-aware energy disaggregation framework [C]// Proceedings of the 2nd ACM International Conference on Embedded Systems for Energy-efficient Built Environments, November 4-5, 2015, Seoul, South Korea: 45-54.

- [26] HEBRAIL G E R, BARARD A E R. Individual household electric power consumption DatSet (IHEPCds) [EB/OL]. [2020-10-21]. <https://archive.ics.uci.edu/ml/datasets/Individual+household+electric+power+consumption>.
- [27] MAKONIN S, POPOWICH F, BARTRAM L, et al. AMPds: a public dataset for load disaggregation and eco-feedback research [C]// 2013 IEEE Electrical Power & Energy Conference, August 21-23, 2013, Halifax, Canada: 1-6.
- [28] RASHID H, SINGH P, SINGH A. I-BLEND, a campus-scale commercial and residential buildings electrical energy dataset[J]. Scientific Data, 2019, 6: 190015.
- [29] 王坤峰, 苟超, 段艳杰, 等. 生成式对抗网络GAN的研究进展与展望[J]. 自动化学报, 2017, 43(3): 321-332.
WANG Kunfeng, GOU Chao, DUAN Yanjie, et al. Generative adversarial networks: the state of the art and beyond [J]. Acta Automatica Sinica, 2017, 43(3): 321-332.
- [30] HE K, ZHANG X, REN S, et al. Deep residual learning for image recognition [C]// IEEE Conference on Computer Vision & Pattern Recognition, June 26-July 1, 2016, Las Vegas, USA: 770-778.
- [31] 张若愚, 吴俊勇, 李宝琴, 等. 基于迁移学习的电力系统暂态稳定自适应预测[J]. 电网技术, 2020, 44(6): 2196-2203.
ZHANG Ruoyu, WU Junyong, LI Baoqin, et al. Self-adaptive power system transient stability prediction based on transfer learning[J]. Power System Technology, 2020, 44(6): 2196-2203.
- [32] KRIZHEVSKY A, SUTSKEVER I, HINTON G. ImageNet classification with deep convolutional neural networks [J]. Advances in Neural Information Processing Systems, 2012, 25(2): 1-6.
- [33] BEE L, SANGHYUN S, HEEWON K, et al. Enhanced deep residual networks for single image super-resolution [C]// Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), July 21-26, 2017, Honolulu, USA: 136-144.
- [34] WU Z, CHEN C W. Signal reconstruction from partial frequency coefficients for image/video frame up sampling [J]. IEEE Transactions on Broadcasting, 2014, 60(3): 575-581.
- [35] CHAI L, SHENG Y. Optimal design of multichannel equalizers for the structural similarity index [J]. IEEE Transactions on Image Processing, 2014, 23(12): 5626-5637.
- [36] CHUNG K L, LIANG Y C, WANG C S. Effective content-aware chroma reconstruction method for screen content images [J]. IEEE Transactions on Image Processing, 2019, 28(3): 1108-1117.
- [37] LIN J. Divergence measures based on the Shannon entropy [J]. IEEE Transactions on Information Theory, 1991, 37(1): 145-151.

李富盛(1996—),男,硕士研究生,主要研究方向:人工智能在电力系统的应用。E-mail:lifusheng0208@foxmail.com

林丹(1996—),女,硕士研究生,主要研究方向:人工智能在电力系统的应用。E-mail:lindan0203@foxmail.com

余涛(1974—),男,通信作者,博士,教授,博士生导师,主要研究方向:复杂电力系统的非线性理论与仿真。E-mail:taoyu1@scut.edu.cn

(编辑 代长振)

Frequency-increased Reconstruction Method for Electrical Data Based on Improved Generative Adversarial Network

LI Fusheng, LIN Dan, YU Tao, WANG Keying, WU Yufeng, YANG Jiajun

(School of Electric Power, South China University of Technology, Guangzhou 510640, China)

Abstract: High-frequency electrical data is one of the data bases to improve the accuracy of situation awareness, monitoring level and auxiliary service quality of power grids. However, the traditional reconstruction algorithm is difficult to achieve the high-precision data reconstruction. Therefore, this paper uses the improved generative adversarial network to reconstruct the low-frequency electrical data into the high-frequency ones. By transforming the time series data into electrical images, the efficient extraction of electrical image features by neural network method is realized. The generator based on deep residual network and improved residual block structure are used to improve the feature learning ability of the generator. In addition, the generator loss function considers the difference in low- or high-dimensional features between real samples and generated samples. The public data set is taken as an example to verify the algorithm. The verification results show that compared with the traditional reconstruction methods, the proposed method has higher peak signal-to-noise ratio, structural similarity, lower mean absolute error and mean absolute percentage error, as well as the higher high-frequency detail reproduction and reconstruction accuracy, and can be generalized for different data sets.

This work is supported by National Natural Science Foundation of China (No. U2066212) and Key Projects of Basic Research and Applied Basic Research in Universities of Guangdong Province (No. 2018KZDXM001).

Key words: frequency-increased reconstruction; low-frequency electrical data; generative adversarial network; deep residual network; image feature; data-driven

