

东莞理工学院

本科毕业论文

毕业设计题目： 基于条件生成对抗网络的 PET 图像重构

学生姓名： 麦伟明

学 号： 201641410119

院 系： 计算机科学与技术学院

专业班级： 2016 级信息与计算科学 1 班

指导教师姓名及职称： 吴志锋（讲师）

起止时间： 2019 年 12 月—— 2020 年 5 月

基于条件生成对抗网络的PET图像重构

麦伟明

2020 年 5 月 12 日

摘 要

正电子放射成像的迭代重建算法一般有两种，一种是统计迭代重建算法一种是代数迭代重建算法，统计迭代算法在PET图像重构领域在近年来得到广泛发展，其中运用最经典的算法为Shepp和Verdi等人的MLEM算法。MLEM算法是基于严格的凸优化理论以及统计理论推导出的算法，其高效且简单。但同时有一个很大缺点，当迭代次数增加，重构过程中的噪声效应也不断加大，到了迭代后期，重构出的图像一般有较高水平的泊松噪声，这一缺陷在低剂量的正电子放射成像里尤其显著。本文根据Kuang Gong与Zhaoheng Xie等人的深度学习去噪方法，对低剂量的MLEM重构结果进行后处理去噪，并应用MAPEM-NET的重构框架对常规剂量的PET数据进行重构。实验结果显示低剂量EM重构的图像经过Unet以及SAGAN去噪后噪声明显减少，图片质量有所提升。在MAPEM-NET中的MAPEM重构由于嵌入了Unet的去噪，使得数据一致性得到保证，图像细节的边缘也能得到恢复。PSNR以及SSIM指标将会用作比较不同方法重构图片质量的评估。

关键词：PET 图像重构、 U型卷积网络、 自注意力机制对抗生成网络、 最大后验正则化

Abstract

The statistical iteration reconstruction and algebraic iteration methods have been widely used in the image reconstruction of Positron Emission Tomography. The MLEM algorithm proposed by Shepp and Verdi is used as a baseline method. This algorithm is based on the convex optimization method and statistical theory, which makes it efficient and reliable. However, the noise of the image will be accumulated continuously in the later iteration with the traditional MLEM algorithm, resulting in the decrease of the image quality. This effect is particularly significant in the low-dose PET image reconstruction. Our work focus on the Deep Learning denoising method in low-dose PET image reconstruction. We applied the Unet and SAGAN denoising methods into the low-dose reconstructed image post-processing, and the MAPEM-NET was used to reconstruct the routine-dose PET data. Experimental results show that Unet and SAGAN denoising can significantly enhance the quality of the reconstructed image. The MAPEN-NET can preserve the data consistency and edge information because of the integrated Unet denoising procession. PSNR and SSIM are adopted to evaluate these two methods and the traditional EM reconstruction algorithm.

Keywords: PET Image Reconstruction, Unet CNN, Self-Attention GAN, Maximum-a-Posterior

目 录

1	背景介绍	5
2	理 论	7
2.1	Poisson分布的特性推导	7
2.1.1	可加性	7
2.1.2	单个随机变量的分布	7
2.2	MLEM算法	8
2.2.1	ECT算法推导	8
2.2.2	E-step	8
2.2.3	M-step	9
2.3	U型卷积神经网络	10
2.3.1	下采样(Down-sampling)	10
2.3.2	上采样(Up-sampling)	11
2.4	自注意力对抗生成网络(SAGAN)	12
3	实验设计	13
3.1	数据集(Data Set)	13
3.2	MAPEM-NET的搭建	13
3.3	SAGAN的搭建	15
4	实验结果	17
4.1	Unet去噪与EM重构	17
4.2	SAGAN去噪与EM重构	17
4.3	SAGAN去噪与Unet去噪	18
4.4	MAPEM-Net重构与EM重构	18
4.5	PSNR与SSIM数值比较	19

4.6 实验总结	20
5 未来工作	21
参考文献	22
致 谢	23

§1 背景介绍

在现代医学中PET作为一种生物医学研究技术和临床诊断手段，其在医学影像中具有极其重要的研究价值。PET成像技术是实现分子水平显像的最佳途径之一，为临床诊断、治疗和预后监测，以及新药研究和开发研制等提供了十分有力的分析手段，是一个极具研究价值的课题。而由于其成像的物理原因是接受放射性物质湮灭释放出来的光子，而不是直接得到我们所需要的信息。因此，如何根据扫描数据来重构出高质量的图像，一直是PET领域的一个非常重要研究课题^[1]。

在PET重建算法中，主要可分为解析法和迭代法两大类。图像重建的解析算法与图像重建的迭代算法之间的主要区别在于对图像的模型化。在解析算法中，一般假设图像像素是连续的。但是在迭代算法中，假设图像像素是离散的。在图像重构中，像素的建模对重建图像的质量好坏影响很大。图像重建的解析算法与图像重建的迭代算法之间的另一个区别是解析算法目标是对一个积分方程求解，而迭代算法则着力对一个线性方程组求解，线性方程组相对于积分方程要容易解，图形重建中迭代算法可以解决更多实际中的成像问题，重建图像也更为准确。

根据正电子成像像素密度服从泊松分布这一性质，Shepp和Verdi^[1]等人提出了基于极大似然估计的MLEM算法，MLEM算法是通过类似于梯度下降方式对目标函数进行优化的迭代算法，由于探测器的重构数据混杂了随机事件以及光子散射和衰减等噪声影响，算法在迭代后期噪声效应会逐渐增大。为了解决这个问题，有人提出在重构目标函数中加入正则项的极大似然估计PML(Penalized Maximum Likelihood)，这种正则项被称作函数的先验，一般来说正则项应该被设计为凸函数，因为这样可以保证能得到其全局最优值。同时，正则项应具有很好的图像细节保护性。近几年来，在正则项上的研究非常多，Wang等人提出基于块像素正则化函数的PML算法^[5]，该算法比传统的基于单个像素点的正则化函数鲁棒性更强，即对边缘信息保护的更加好。Nutys^[6]等人对正则化函数里像素间的距离度量方式进行优化，提出一种基于相关差异性(Relative Difference)的正则化函数，实验结果验证该方法能避免图像边缘信息的过度平滑化。

以上是对于PET重构模型上做出的改进，同时也有许多学者对目标函数的给出相应的优化方法，如Chun等提出的ADMM算法优化^[7]带正则项的目标函数，相对于其他方法如EM、OSEM、梯度下降方法等，其收敛速度更加快。有了这些相关的技术基础，Kuang Gong等学者提出了运用深度学习^[9]方法融入MAPEM重构目标函数的优化过程中，在ADMM的迭代过程中加入Unet卷积网络以及残差网络，构建了MAPEM-NET的算法结构，Unet能够有效地降低EM重构过程中的噪声传播同时也能减少细节信息的损失。由于Unet的去噪过程与MAPEM的更新过程融合在一起，因此重构图像数据的一致性能得到保证。Jelmer M. 等在2017提出GAN网络对低剂量的PET图像进行重构^[8]，经过GAN的处理，低剂量PET重构像素密度能够达到常规剂量图像级别，意味着重构出的图像清晰度更加高。受到Kuang Gong的启发，Zhaoheng Xie团队提出将SAGAN(Self-Attention Generative Adversarial Network) 替换Unet，在训练数据集有限的情况下，GAN 的去噪效果比Unet 去噪效果要好，而且因为加入了自注意力机制，图像的纹理也能得到较好的恢复。

本次毕业设计的工作，主要是复现MLEM重构算法，以及学习卷积神经网络、对抗生成网络的理论内容。理解Kuang Gong等学者在PET图像重构方面的研究工作，对其中的基于深度学习算

法的PET图像重构进行复现，对数值结果进行简单对比。在后面几个小节里会这样安排，第二节是MLEM算法的统计理论基础以及Unet和SAGAN涉及的深度学习知识，第三节是具体实验设计以及算法流程，具体数据集上的运行情况。第四节是实验结果的呈现以及对比。

§2 理 论

由于基于统计迭代的PET图像重构算法是根据正电子放射符合泊松分布这个特性来设计的，因此在此具体介绍MLEM算法之前会对泊松分布以及泊松过程的一些特性进行探究，以下先对泊松分布以及EM 算法理论基础进行介绍，然后是Lange等提出^[10] 的Emission Computed Tomography 算法的推导。最后介绍实验中的Unet 以及SAGAN的结构的设计及其中涉及理论。

§2.1 Poisson分布的特性推导

§2.1.1 可加性

若非负随机变量 Z_i 服从泊松分布， $Z_1 \sim P(\lambda_1)$ ， $Z_2 \sim P(\lambda_2)$ ， $Z = Z_1 + Z_2$ ，则 $Z \sim P(\lambda)$ ， $\lambda = \lambda_1 + \lambda_2$ 。简单证明：

$$\begin{aligned} P\{Z = k\} &= P\{Z_1 + Z_2 = k\} \\ &= \sum_{m=0}^k P\{Z_1 = m\}P\{Z_2 = k - m\} \\ &= \sum_{m=0}^k \frac{e^{-\lambda_1} \lambda_1^m}{m!} \frac{e^{-\lambda_2} \lambda_2^{k-m}}{(k-m)!} \\ &= \frac{e^{-(\lambda_1+\lambda_2)}}{k!} \sum_{m=0}^k \frac{k!}{m!(k-m)!} \lambda_1^m \lambda_2^{k-m} \end{aligned}$$

§2.1.2 单个随机变量的分布

若 $Z = \sum_{k=1}^m Z_k$ ，随机变量 Z_k 出现的概率为 p_k ，在随机变量 Z 服从 $P(\lambda)$ 的泊松分布的情况下，以下成立：

$$Z_k \sim P(\lambda p_k) \quad (2.1)$$

证明如下：

$$\begin{aligned} P\{Z_k = m\} &= \sum_{n=0}^{m+n} P\{Z_k = m, Z = m + n\} \\ &= \sum_{n=0}^{m+n} P\{Z_k = m | Z = m + n\} P\{Z = m + n\} \\ &= \sum_{n=0}^{m+n} \binom{m}{m+n} p_k^m (1 - p_k)^n \frac{e^{-\lambda} \lambda^{m+n}}{(m+n)!} \\ &= e^{-\lambda} \sum_{n=0}^{m+n} \frac{(\lambda p_k)^m (1 - p_k)^n \lambda^n}{m! n!} \\ &= e^{-\lambda p_k} \frac{(\lambda p_k)^m}{m!} \end{aligned}$$

§2.2 MLEM算法

最大化期望算法(Expectation Maximization), 是一类通过迭代进行极大似然估计的优化算法, 1977年由Dempster等人总结提出, 用于含有隐变量(latent variable)的概率模型参数的极大似然估计, 或最大后验概率估计, EM算法每次迭代由两部分交替完成^[12]: E步, 求期望(Expectation); M步, 求极大(Maximization). 由于迭代规则容易实现并考虑到隐变量, EM算法被广泛应用于处理数据的缺失值, 以及很多机器学习算法, 包括高斯混合模型(Gaussian Mixture Model, GMM) 和隐马尔可夫模型(Hidden Markov Model, HMM)的参数估计。

§2.2.1 ECT算法推导

正电子放射成像的目标是要估计成像区域中人体组织处的光子密度, 从组织处放射的光子收到各种影响会偏离原来的发射轨迹^[4], 也会有些光子在放射过程中因为能量的衰减而无法到达探测器, 这些因素必须考虑在重构算法里。将EM算法应用在放射图像重构里的方法称作ECT算法(Emission Computed Tomography)。在PET的数据建模中, 探测器探测到的数据 $\mathbf{y} \in \mathbb{R}^{M \times 1}$ 与重构图像 $\mathbf{x} \in \mathbb{R}^{N \times 1}$ 相关, \mathbf{M} 是探测器个数, \mathbf{N} 是重构图像像素数。设光子的散射事件以及随机事件为 $\mathbf{s} \in \mathbb{R}^{M \times 1}$ 和 $\mathbf{r} \in \mathbb{R}^{M \times 1}$, 系统矩阵为 $\mathbf{P} \in \mathbb{R}^{M \times N}$, 那么探测数据与重构图像有如下关系:

$$\mathbf{y} = \mathbf{P}\mathbf{x} + \mathbf{s} + \mathbf{r} \quad (2.2)$$

其中系统矩阵中的元素 P_{ij} 代表重构区域中每个像素点 j 被第 i 个探测器检测到的概率^[9]。因此重构的目标似然函数可写为:

$$L(\mathbf{y}|\mathbf{x}) = \sum_{i=1}^M y_i \log \bar{y}_i - \bar{y}_i - \log y_i! \quad (2.3)$$

因此, 要得到我们最终的图像数据 $\hat{\mathbf{x}}$ 需要解如下的优化问题:

$$\hat{\mathbf{x}} = \arg \max_{\mathbf{x} \geq 0} L(\mathbf{y}|\mathbf{x}) \quad (2.4)$$

至此, 我们EM算法优化的目标函数就已经建立好, 以下是具体EM算法处理PET 图像重构的过程推导。

§2.2.2 E-step

EM算法的第一步操作就是求似然函数的期望, 首先根据极大似然估计方法给出数学模型:

$$\begin{aligned} \max_{\lambda} L(\lambda) &= \log \prod_{i=1}^N p\{Y = y_i | \lambda\} \\ &= \log \prod_{i=1}^N \sum_{\mathbf{x}_i \in I_i} p\{Y = y_i, X_i = \mathbf{x}_i | \lambda\} \end{aligned} \quad (2.5)$$

这里 λ 代表重构图像里每个像素点所服从的泊松分布的参数，亦即药剂强度。 $\mathbf{X}_i = (x_1^i, x_2^i, \dots, x_j^i)$ 为随机向量，每一维代表每个像素点发生湮灭所产生的粒子数。因为 $y_i = \sum_j x_j^i$ ，所以有：

$$\log \prod_{i=1}^N \sum_{\mathbf{x}_i \in I_i} p\{Y = y_i, X_i = x_i | \lambda\} = \log \prod_{i=1}^N \sum_{\mathbf{x}_i \in I_i} p\{X_i = x_i | \lambda\} \quad (2.6)$$

设 $Q(x^i) = P(X^i | Y_i, \lambda)$ ， $\sum_{\mathbf{x}_i \in I_i} Q(x^i) = 1$ ，由此得出带有隐变量的似然函数：

$$\max_{\lambda} L(\lambda) = \sum_{i=1}^k \log \sum_{\mathbf{x}_i \in I_i} \frac{Q(x_i) p\{X_i = x_i | \lambda\}}{Q(x^i)} \quad (2.7)$$

由Jensen不等式可得：

$$\begin{aligned} L(\lambda) &\geq \sum_{i=1}^k \sum_{\mathbf{x}_i \in I_i} Q(x^i) \log p\{X_i = x_i | \lambda\} \\ &= \sum_{i=1}^k \sum_{\mathbf{x}_i \in I_i} P(X^i | Y_i, \lambda) \log p\{X_i = x_i | \lambda\} \\ &= \sum_{i=1}^k E_{x^i} (\log p(X^i | \lambda) | Y^i, \lambda) \\ &= \sum_{i=1}^k \sum_j E_{x^i} (\log p(X_j^i | \lambda) | Y^i, \lambda) \end{aligned} \quad (2.8)$$

x_j^i 代表 x^i 里的某个分量，即某个像素点的强度值，简单地说我们通过Jensen不等式可以得到似然函数的下界，只要通过不断最大化其下界，那么似然函数也可以达到最大化。到此EM算法的E步完成，根据以上公式可计算出当前参数下概率的期望值。然后是M步，此步骤较简单，就是纯粹的求导，令导函数为0求驻点。

§2.2.3 M-step

$$E_{x^i} (\log (X_j^i | \lambda) | Y_i, \lambda) = \sum_{x_i \in I_i} \log P(X_j^i | Y_i, \lambda) P(X_j^i | \lambda) \quad (2.9)$$

由于 $X_j^i \sim P(c_{ij}\lambda_j)$ ， c_{ij} 为衰减系数与系统矩阵概率作用后相应光子到达探测器的概率，式2.9可转化为如下形式：

$$\begin{aligned} \sum_{x_i \in I_i} P(X_j^i | Y_i, \lambda) - c_{ij}\lambda_j + X_j^i \log(c_{ij}\lambda_j) - \log(X_j^i) \\ = -c_{ij}\lambda_j + E(X_j^i | Y_i, \lambda) \log(c_{ij}\lambda_j) - \log(X_j^i) \end{aligned} \quad (2.10)$$

因为 X_j^i 在 Y_i , λ 条件下服从多项式分布, 由已有结论可知2.10中的期望可如下计算:

$$E(X_j^i|Y_i, \lambda) = \frac{c_{ij}\lambda_j Y_i}{\sum_{k \in I_i} c_{ik}\lambda_k^n} \quad (2.11)$$

对以上求导令导数为0, 可得第 $n+1$ 次迭代第 j 像素点的强度值的估计:

$$\lambda_j^{n+1} = \frac{\lambda_j^n}{\sum_{i \in I_i} c_{ij}} \sum_{i \in I_i} \frac{c_{ij} Y_i}{\sum_{k \in I_i} c_{ik} \lambda_k^n} \quad (2.12)$$

§2.3 U型卷积神经网络

Unet是卷积神经网络的一种, 2015年由Olaf 等学者提出, 但其网络结构与经典的VGG卷积网络不同, 此网络在医学图像分割任务上有出色的表现。Unet可分为两个主体, 一个是encoder, 一个是decoder。Encoder里面主要的操作就是卷积层的卷积和池化层的池化, decoder的操作是上采样以及和encoder里每个卷积层的feature map进行合并。传统Unet是用来对图像进行分割, 可以应用在医学图像上, 将病变组织和其他正常组织分割开来。

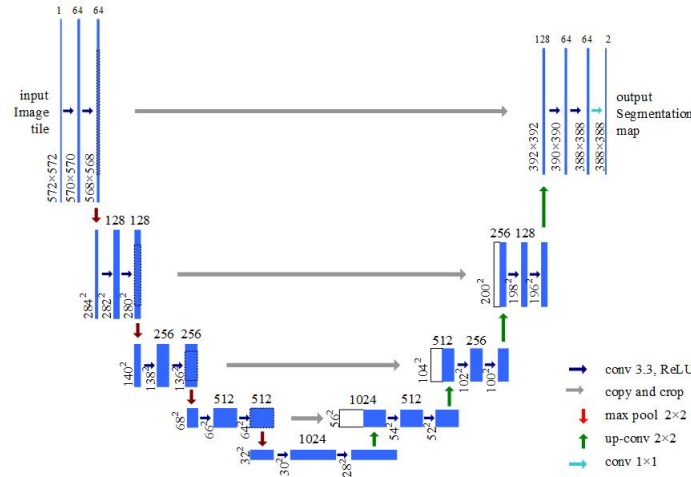


图 1: Olaf的U型神经网络

上图展示的是Olaf等人提出的U网络结构图, 但在Kuang Gong等人的论文里, 他们将原来U网络的结构进行修改, 以达到去噪效果以及减少参数的目的, 下面具体介绍二者的不同之处。

§2.3.1 下采样(Down-sampling)

Unet的前半部分做的是特征提取的工作, 主要有卷积层(Convolution Layer)以及最大化池化层(Max Pooling), 最大池化层的作用一般来说是降维, 减少网络的学习参数, 防止过拟合。而在Kung Gong 的Unet里面, 池化层变为了 $strides = 2$ 的卷积层, 因此他们的Unet是只有卷积层的。根据卷积

核(filter)的大小以及卷积的步长strides,可以由以下公式确定卷积处理后的图像大小:

$$N = \frac{W - F + 2P}{S + 1} \quad (2.13)$$

其中, N 为卷积后图像宽大小, W 为输入图片大小, F 为filter的大小, S 为strides步长, P 是padding补零的数目, padding有两种模式, 一种是valid, 一种是same, valid的padding不会在原图像边缘上补零, same模式会在原图补零使得原图像与卷积后图像大小一致。

§2.3.2 上采样(Up-sampling)

上采样是由反卷积的方式完成的, 反卷积(Deconvolution)的概念第一次出现是Zeiler在2010年发表的论文Deconvolutional networks中, 但是并没有指定反卷积这个名字, 反卷积这个术语正式的使用是在其之后的工作中(Adaptive deconvolutional networks for mid and high level feature learning)。随着反卷积在神经网络可视化上的成功应用, 其被越来越多的工作所采纳比如: 场景分割、生成模型等。其中反卷积(Deconvolution)也有很多其他的叫法, 比如: Transposed Convolution, Fractional Strided Convolution等等。

通俗来讲, 反卷积是通过对降维后图像进行卷积操作, 令其还原为原来未降维时的大小, 过程与卷积的操作一样, 只是在输入图像的边缘会补上足够的0。在Kuang Gong的论文里, 反向卷积后的feature map会与下采样过程中的feature map相加而不是叠在一起, 因此上采样后的feature map高度并不会增加, 这样可以起到减少网络参数的作用。最后输入图像会与最后一个卷积层的输出相加, 形成一个残差网络。

64

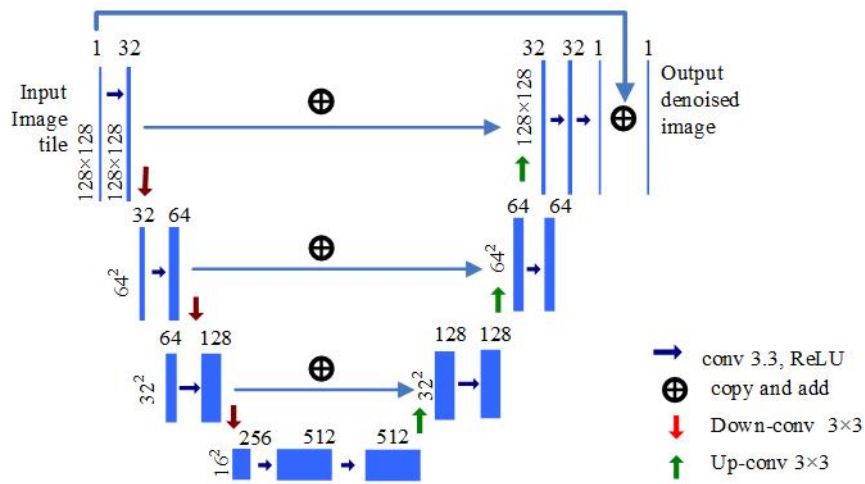


图 2: Kuang Gong的U型神经网络

§2.4 自注意力对抗生成网络(SAGAN)

Ian J. Goodfellow等人于2014年提出了一个通过对抗过程估计生成模型的新框架。框架中同时训练两个模型：捕获数据分布的生成模型(Generator)，和估计样本来自训练数据的概率的判别模型(Discriminator)。生成器的目标是在判别器指引下完成某个特定任务，而且该任务的输出是要被判别器所接受的。以下将生成器称作 G ，判别器称作 D ，这个框架对应一个最大值集下限的双方对抗游戏。可以证明在任意函数 G 和 D 的空间中，存在唯一的解决方案，使得 G 重现训练数据分布，使得判别器 D 无法判别生成器输出的真伪。在 G 和 D 由多层感知器定义的情况下，整个系统可以用反向传播进行训练。在训练或生成样本期间，不需要任何马尔科夫链或展开的近似推理网络。实验通过对生成的样品的定性和定量评估证明了本框架的潜力，这种框架也就是对抗生成网络。

由于传统GAN在图像细节方面处理得不够好，Goodfellow团队在2018年提出基于自注意力机制的对抗生成网络(SAGAN)。在SAGAN中，生成器能够通过特征的位置信息对图像的细节还原^[13]。另外，判别器能够检测到图像中的不同纹理的连续性。这些优点都是由于SAGAN运用了自注意力机制(Self-Attention)，该机制是*Nonlocal Neural Networks*这篇文章提出的。

在卷积神经网络中，每个卷积核的尺寸都是有限的，因此每次卷积操作只能覆盖像素点周围很小一块邻域。对于距离较远的特征，例如狗有四条腿这类特征，就不容易捕获到，池化操作或许能够让网络探测到距离较远的特征，它会把feature map的高和宽变得越来越小，越靠后的层，其卷积核覆盖的区域映射回原图对应的面积越大。但是毕竟还得需要经过多层映射，不够直接。Self-Attention通过直接计算图像中任意两个像素点之间的相关性，一步到位地获取图像的全局几何特征。

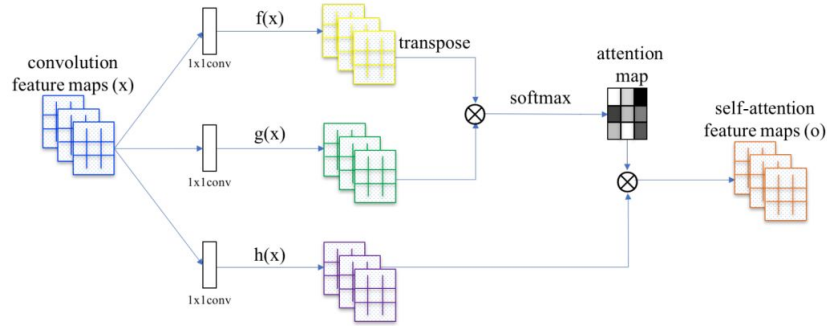


图 3: 将自注意力机制运用在GAN中的方式。符号 \otimes 代表矩阵对应元素相乘，softmax函数被用作激活函数

图3中有三个函数操作，经 $g(x)$ 与 $f(x)$ 处理后的feature map会进行resize处理，并且相乘形成attention map。从某种意义上讲，attention map和相关性矩阵非常相似，它描述的是每个像素点与其他位置像素的相关性。

§3 实验设计

在这一节里，主要介绍我在本毕设项目里的实验细节，相关算法的伪代码，所有实验均在Matlab与Python环境下实现。

§3.1 数据集(Data Set)

本项目的实验数据来源于BrainWeb^[14]，BrainWeb是脑部的模拟数据库(Simulated Brain Databasse)，SBD包含两种解剖模型的模拟脑部MRI数据：正常脑部以及多发性硬化症脑部。这些脑部三维数据用三种序列进行模拟(T1-, T2-, and proton-density-(PD-) weighted)，每个模型有不同的切片厚度，噪声级别以及强度级别。实验中选用了其中一个脑部模型，共有181张slices，选取其中19张作为Unet的训练集，数据集格式为.rawb文件，此文件共有三维 $217 \times 181 \times 181$ ，前两维为图片长和宽，第三维为不同高度下的slice。利用Matlab对.rawb图像文件进行处理将181帧图像分割开来，另存为png图片。ground truth 图片和噪声图片分别存放在三个文件夹。由于神经网络输入图片大长宽相同时，卷积操作比较方便，因此对原数据进行resize操作， 217×181 转为 128×128 。所用搭建的神经网络API为Keras，数据输入维度为4，因此用reshape函数对其维度进行增加。

得到图像数据后，放入Matlab的模拟程序中生成不同剂量级别的正电子重构探测数据sinogram data，在这里我们用光子数为 $1.36E6$ 数据作为低剂量重构数据(Low-count)，光子数为 $1.36E7$ 为高剂量重构数据(High-count)^[9]，数量为low count数据的10倍，然后以low count数据用EM算法重构50步得到噪声图像为Unet的输入，high count的EM算法重构图像作为标签。

§3.2 MAPEM-NET的搭建

这里根据Kuang Gong^[2]的论文，对传统Unet进行改动，池化层改为步长为2的卷积，这个替换同样可以压缩图像大小为原来二分之一。Encoder 和Decoder 里的卷积层数量有所减少，原先的feature map的合并改为相加，减少参数的更新。最后输出为网络输入加上最后一层卷积层的输出。这里训练过程还有一点与传统Unet训练不同的是，训练集采取带噪声的低剂量切片重构图像，标签为高剂量的切片重构图像，而传统Unet的标签是分割的图片。不能够直接使用ground truth作为标签训练，因为ground truth数据格式为.rawb，像素取值范围在 $0 \sim 255$ 之间，直接用此数据训练在MAPEM-NET里会造成数据不一致的问题。

将训练后的Unet加入到带正则项的重构目标函数里，用ADMM的算法迭代优化目标函数，就是Kuang Gong等人的MAPEM-NET框架。以下是MAPEM-NET的优化迭代函数：

$$\max_{\mathbf{x}, \mathbf{z}} L(\mathbf{y}|\mathbf{x}) - \beta U(\mathbf{z}), \text{ s.t. } \mathbf{x} = \mathbf{z} \quad (3.1)$$

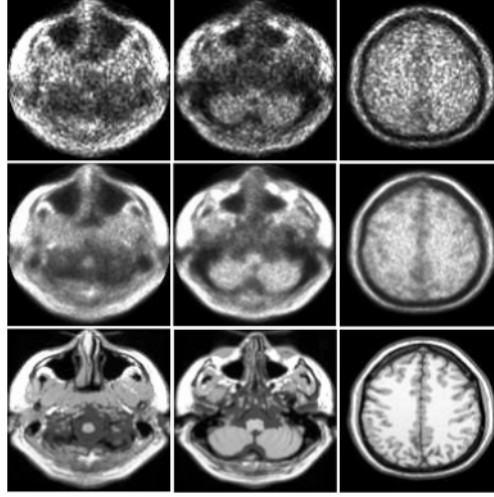


图 4: 训练集: 其中三个脑部切片, 第一行为low count数据即神经网络输入, 第二行为high count数据训练标签, 第三行为ground truth

基于ADMM的优化算法有如下迭代格式:

$$\begin{aligned}
 \mathbf{x}^{n+1} &= \arg \max L(\mathbf{y}|\mathbf{x}) - \rho \|\mathbf{x} - \mathbf{z}^n + \boldsymbol{\mu}^n\|_2 \\
 \mathbf{z}^{n+1} &= \arg \min \|\mathbf{x}^{n+1} - \mathbf{z} + \boldsymbol{\mu}^n\|_2 + \beta U(\mathbf{z}) \\
 \boldsymbol{\mu}^{n+1} &= \boldsymbol{\mu}^n + \mathbf{x}^{n+1} - \mathbf{z}^{n+1}
 \end{aligned} \tag{3.2}$$

式3.2中, \mathbf{x} 为重构图像, 对 \mathbf{z} 的更新将由 $\mathbf{z}^{n+1} = f(\mathbf{x}^{n+1})$ 替代, $f(\mathbf{x})$ 为Unet处理后的图像。

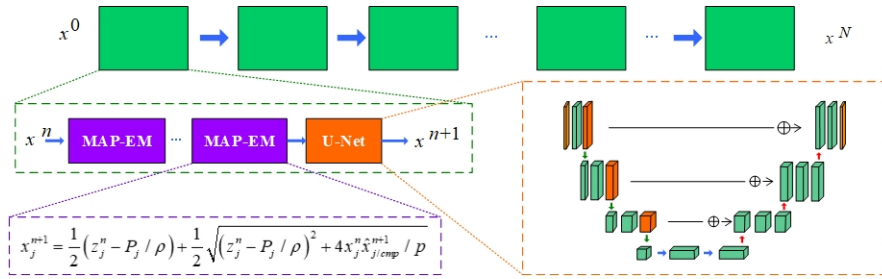


图 5: MAPEM-NET框架

Algorithm 1 MAPEM-NET for PET Image Reconstruction

Input: Number of Modules $MaxIt$, sub-iteration number $SubIt$
Output: Reconstructed Image z

```

1: Initialize  $x^0, z^0, \mu^0, \rho$ 
2: repeat
3:   for  $n = 1$  to  $SubIt$  do
4:      $\hat{x}_{temp}^{n+1} = P^T \frac{y}{P x^n + r}$ 
5:      $x_j^{n+1} = \frac{1}{2} [z_j^n - \mu_j^n - P_{.j}/\rho] + \frac{1}{2} \sqrt{[z_j^n - \mu_j^n - P_{.j}/\rho]^2 + 4x_j^n \hat{x}_{j,temp}^{n+1} / \rho}$ 
6:      $\mu^{n+1} = \mu^n + x_j^{n+1} - \hat{x}_{temp}^{n+1}$ 
7:   end for
8:    $z^{n+1} = f(x^{n+1})$ 
9:    $\mu^{n+1} = \mu^n + x^{n+1} - z^{n+1}$ 
10: until Maximum Modules
11: return  $z$ ;
    
```

根据Kuang Gong等的算法稍作改动，在内循环中只更新 x 和 x_{temp} ，然后再交给Unet进行去噪处理。

§3.3 SAGAN的搭建

SAGAN的实现，是根据Zhaoheng Xie^[3]等人论文的方法进行搭建，生成器为改进后的Unet，更加适用于去噪工作，判别器为VGG型卷积神经网络，最后的输出激活函数为 $sigmoid$ 。其具体结构图如下：如之前所说SAGAN与传统GAN有所区别，其中的Self-Attention机制通过直接计算图像中任

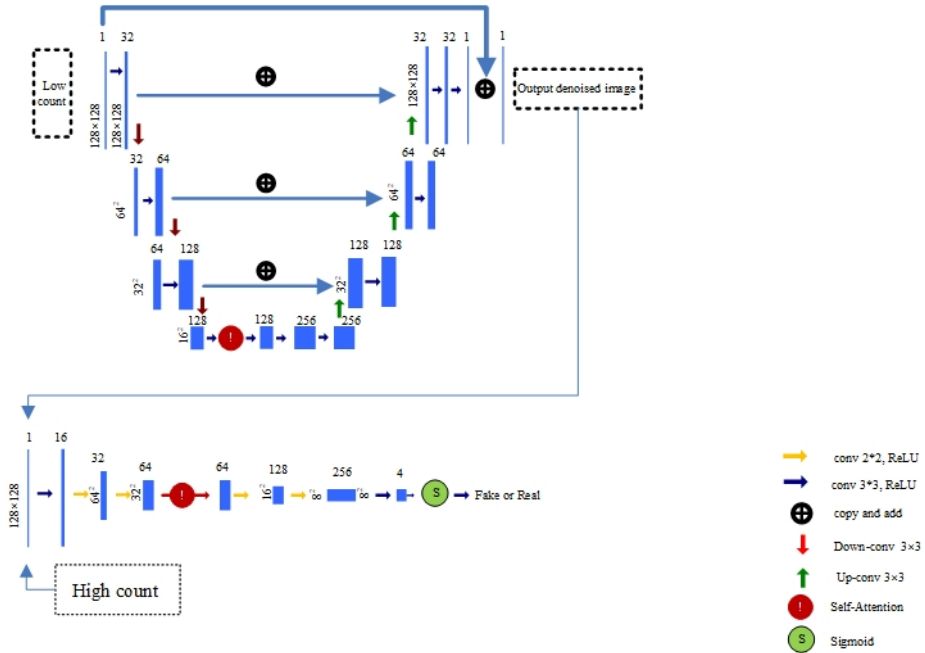


图 6: SAGAN框架

意两个像素点之间的关系，一步到位地获取图像的全局几何特征。它的优化方式与传统GAN一样，

先更新判别器，然后固定住判别器的权值，更新生成器。以下是本次实验所采取的优化损失函数：

$$\begin{aligned} L_G &= \lambda_1 \|G(\boldsymbol{\alpha}) - \boldsymbol{x}\|^2 + \lambda_2 l_{bce}(D(G(\boldsymbol{\alpha})), 1) \\ L_D &= l_{bce}(D(\boldsymbol{x}), 1) + l_{bce}(D(G(\boldsymbol{\alpha})), 0) \end{aligned} \tag{3.3}$$

l_{bce} 为二进制交叉熵函数。首先在high count中训练判别器判别为1，再在生成器生成的数据中训练判别器判别为0，然后训练生成器。循环至对生成器的输出判别器无法判别为0。

§4 实验结果

§4.1 Unet去噪与EM重构

这个小节主要呈现Unet对EM重构后图像进行去噪的效果，所选的数据为low count光子数图像，EM算法重构步数为50步。可以看到，经过Unet处理后的图像噪声明显下降，但在细节复原

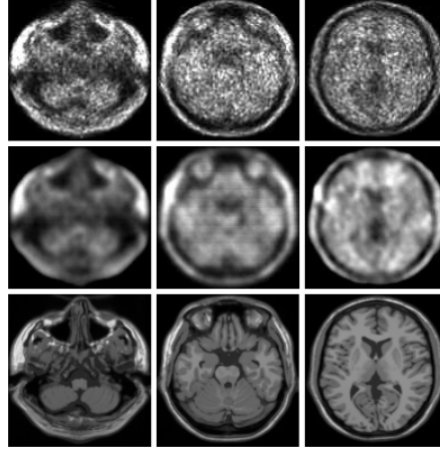


图 7: Low count测试集: 第一行为EM算法重构50步结果, 第二行为经过Unet后处理的EM重构结果, 最后一行为ground truth

方面效果不尽人意, 同时可以看到Unet处理后的图像有些许横向断层, 猜测这是由于数据一致性不够导致, 后续或许可以通过其他方法, 如WTV(Weighted Total Variation)等, 加强图像数据一致性。

§4.2 SAGAN去噪与EM重构

以下是SAGAN对low count数据后处理的结果, 下面放大图像具体部分细节, 可以看到SAGAN细

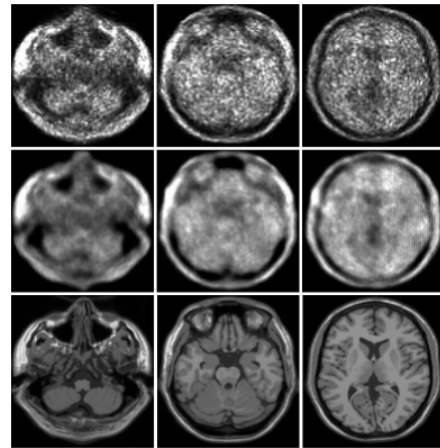


图 8: Low count测试集: 第一行为EM算法重构50步结果, 第二行为经过SAGAN后处理的EM重构结果, 最后一行为ground truth

节复原度更加强, 在关键部位, EM算法重构由于噪声太大导致变形, 经过SAGAN处理后形状可以

得到恢复。而目标处的边缘更加清晰。

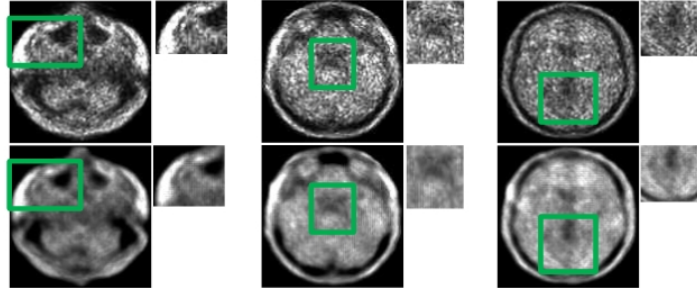


图 9: SAGAN与EM算法关键部位细节对比

§4.3 SAGAN去噪与Unet去噪

下面是Unet后处理效果与SAGAN后处理效果图对比，第三行图片为high count EM重构图像作为参照。从局部细节看，Unet与SAGAN的去噪效果都是非常好，与未处理之前的EM重构相比图像变得更加平滑，但是由于图像变得平滑，Unet处理后的图像变得更加模糊，边缘不够清晰，同时由于数据不一致性导致横向的artifacts。SAGAN能够还原边缘信息，因此看起来比Unet稍清晰。

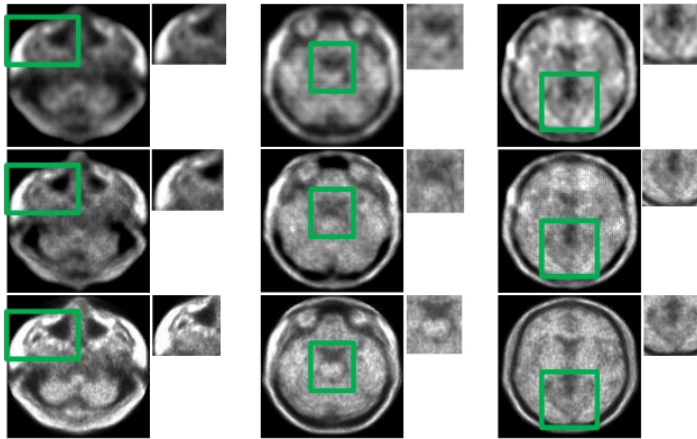


图 10: Unet与SAGAN关键部位细节对比

§4.4 MAPEM-Net重构与EM重构

在MAPEM-NET实验里，采用的重构数据光子数为Routine count，即 $6.8E6$ 光子数，以下给出经过MAPEM-NET处理以及单纯Unet处理的实验结果。这里MAPEM-NET的Modules共5个，即有5个Unet参与处理过程，内循环迭代40次。

图11中第一二行均为光子数 $6.8E6$ 的数据的重构图像，可以看到MAPEM-NET的输出不但噪声程度更加低，细节还原度高，边缘更加分明易于辨认。其效果达到甚至超过第三行High count EM重构的图像的还原程度。

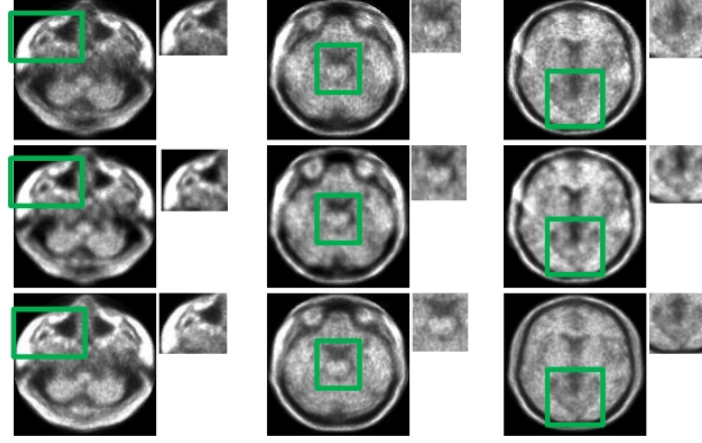


图 11: Routine count光子数测试集：第一行为routine count数据EM算法重构50步，第二行为routine count数据MAPEM-NET算法重构结果，第三行为high count数据EM50步重构结果

§4.5 PSNR与SSIM数值比较

PSNR(Peak Signal to Noise Ratio)即峰值信噪比，是一种评价图像的客观标准，其计算公式如下：

$$PSNR = 10 \times \log_{10} \left(\frac{(2^n - 1)^2}{MSE} \right) \quad (4.1)$$

其中 MSE 为两幅图像的均方误差。一般来讲PSNR值越高越好，但PSNR值的评价有局限性，其数值表现往往和人眼视觉有所差异，PSNR低不一定意味着人眼看上去图片效果更差。

SSIM(Structural Similarity Index)，结构相似性，是一种衡量两幅图像相似度的指标。结构相似度指数从图像组成的角度将结构信息定义为独立于亮度、对比度的，反映场景中物体结构的属性，并将失真建模为亮度、对比度和结构三个不同因素的组合。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度量。其计算公式如下：

$$SSIM(x, y) = \frac{(2\mu_x\mu_y + c_1)(2\sigma_{xy} + c_2)}{(\mu_x^2 + \mu_y^2 + c_1)(\sigma_x^2 + \sigma_y^2 + c_2)} \quad (4.2)$$

x, y 是待比较的两幅图像， σ 为图像标准差， μ 为图像均值。 c_1, c_2 是维持计算稳定的常数，SSIM的值介于0~1，越接近1图像越相似。以下是三幅测试图像的PSNR 值以及SSIM值。

表 1: 第一份测试数据

	EM	U-net	SAGAN
PSNR	22.01	26.73	26.46
SSIM	0.59	0.70	0.68

从表格的数据可看出加了网络去噪的算法比原EM重构算法重构出的图像质量更加高，SAGAN比U-net的PSNR值稍好，但并不突出。而SSIM值中SAGAN的值普遍比Unet低，猜测是GAN中的生成器

表 2: 第二份测试数据

	EM	U-net	SAGAN
<i>PSNR</i>	25.52	26.87	27.06
<i>SSIM</i>	0.56	0.65	0.63

表 3: 第三份测试数据

	EM	U-net	SAGAN
<i>PSNR</i>	23.19	22.04	23.89
<i>SSIM</i>	0.55	0.66	0.63

在还原图像时虚构了一些图像结构细节。在测试量少的情况下结论很难给得出，但从肉眼观测上看SAGAN方法确实比U-net方法清晰度更高。

§4.6 实验总结

本次实验共做了四组对比，前三组分别是U-net、SAGAN与EM重构的对比，和U-net处理效果与SAGAN处理效果的对比。U-net网络对于图像去噪效果非常明显，但是由于对图像过度平滑的处理使得图片看起来非常模糊，而SAGAN由于生成器具有一定的细节还原能力，并且加入了自注意力机制，因此对于图像边缘处理更加到位。

最后一组实验是在Routine count数据中完成的重构，可明显看到原EM算法重构的图像噪声水平依然不低，而MAPEM-NET的处理过程中加入了U-net去噪，以及EM的重构迭代，因此对细节复原度很高，目标边缘更加深色。

§5 未来工作

低剂量PET对病人身体损害更加低，传统EM算法重构的噪声过大影响医疗诊断，通过MAPEM-Net算法以及SAGAN算法进行的PET图像重构结果更加清晰，它们在PET图像重构上有很大的研究前景。

本次实验简单地采取数值指标对重构图片进行分析，因此不能够给出应用深度学习方法的重构比传统MLEM重构效果好这个结论，但是仅从肉眼观测上看，使用了Unet，MAPEM-NET或者SAGAN的重构图像确实比普通EM算法重构的图像边缘信息保护的更加好。后续可以采取相应的数值指标方法对这三种方法的重构结果进行大规模测量，例如PSNR(Peak Signal to Noise Ratio),或者是Kuang Gong论文中提到的CRC(Contrast Recovery Coefficient)指标。

参考文献

- [1] Shepp L A , Vardi Y P . Maximum Likelihood Reconstruction for Emission Tomography[J]. IEEE Transactions on Medical Imaging, 1982, 1(2):113-122.
- [2] Gong K , Wu D , Kim K , et al. MAPEM-Net: an unrolled neural network for Fully 3D PET image reconstruction[C]. The FIM on Fully Three-dimensional Image Reconstruction in Radiology. 2019.
- [3] Xie Z , Baikejiang R , Gong K , et al. Generative adversarial networks based regularized image reconstruction for PET[C]. The FIM on Fully Three-Dimensional Image Reconstruction in Radiology and Nuclear Medicine. 2019.
- [4] Bailey D L , Townsend D W , Valk P E . Positron Emission Tomography[M]. Springer London, 2013.
- [5] Wang G , Qi J . Penalized Likelihood PET Image Reconstruction Using Patch-Based Edge-Preserving Regularization[J]. IEEE Transactions on Medical Imaging, 2012, 31(12):2194-2204.
- [6] Nuyts J , Beque D , Dupont P , et al. A concave prior penalizing relative differences for maximum-a-posteriori reconstruction in emission tomography[C]// Nuclear Science Symposium Conference Record. IEEE, 2002.
- [7] Chun S Y , Dewaraja Y K , Fessler J A . Alternating Direction Method of Multiplier for Tomography With Nonlocal Regularizers[J]. IEEE Transactions on Medical Imaging, 2014, 33(10):1960-1968.
- [8] Wolterink J M , Leiner T , Viergever M A , et al. Generative Adversarial Networks for Noise Reduction in Low-Dose CT[J]. IEEE Transactions on Medical Imaging, 2017, 36(12):2536-2545.
- [9] Gong K , Guan J , Kim K , et al. Iterative PET image reconstruction using convolutional neural network representation[J]. IEEE transactions on medical imaging, 2018, 38(3): 675-685.
- [10] Lange K , Carson R . EM Reconstruction Algorithms for Emission and Transmission Tomography[J]. Journal of Computer Assisted Tomography, 1984, 8(2):306-316.
- [11] Gong T Z. Positron Imaging Reconstruction Technology [D]. Zhejiang University, 2001.
- [12] Hang L. Statistical Learning Methods [M]. Tsinghua University Press, 2012.
- [13] Zhang H, Goodfellow I, Metaxas D, et al. Self-attention generative adversarial networks[J]. arXiv preprint arXiv:1805.08318, 2018.
- [14] C.A.Cocosco,V.Kollokian,R.K.S.Kwan,et al. “Brainweb:Online inter facetoa 3D MRI simulated brain database,” NeuroImage (1997).

致 谢

致我的家人、朋友、我大学里遇到的老师以及所有的经历。