

CVPR 2018 Paper: An Unsupervised Learning Model for
Deformable Medical Image Registration

Guha Balakrishnan et al. MIT

11821011 侯宇轩

分类：医学影像

一、摘要

本文设计了一种新的非监督的三维神经网络用于医学图像配准。方法特色如下：

1) 快速。一般传统配准方法要使用迭代法进行优化，对单个 CT/MRI 数据配准时间可能达到 20 分钟以上。本文网络训练完成后，可以快速进行配准。

2) 非监督。只需要将不同期图像直接输入作为训练数据，网络会自行学习图像之间的对应关系，而不需要人工标注的点对作为训练数据。

二、背景介绍

图像配准是医学人工智能中的难题。

配准指找出一个固定图像与一个变换图像之间的对应关系，并将变换图像向尽量与固定图像相似的方向转变。

医学图像配准主要解决对相同组织/部位的不同次扫描数据的对应问题。一般 CT 扫描要向病人血管中注射造影剂以便显像，随时间造影剂在体内的分布不同，CT 图像中增强（高亮）的部分也不同。以肝脏为例，一般分为动脉期(Arterial)、门静脉期(Portal)和延迟期(Delayed)三期图像。

为什么需要配准？以肝脏肿瘤分割为例，一般我们使用肿瘤边界最为清晰的门静脉期作为训练数据。然而动脉期、延迟期相比门静脉期蕴含了额外的可用信息，如果能够利用它们，那么在加大数据量、提高精度等等方面有非常大的潜力。

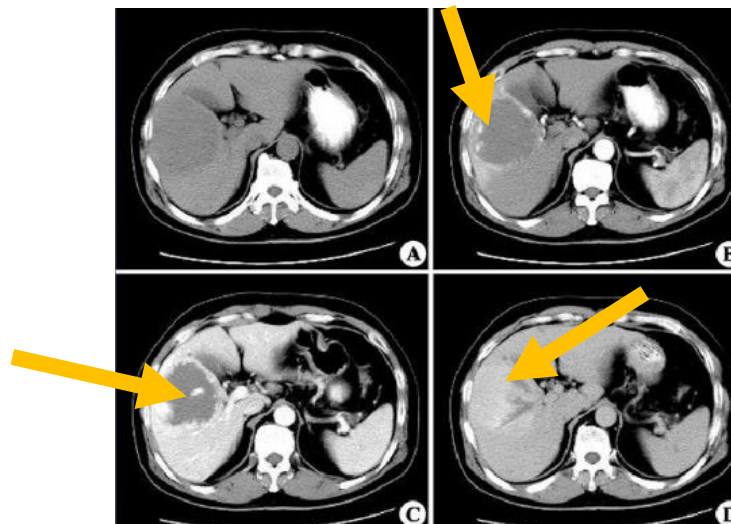


图 1：肝脏多期 CT 图像。A:平扫。B:动脉期。C:门静脉期。D:延迟期。可以明显的看出，不同期肝脏图像同一位置（箭头所指为肿瘤）强度不同。

三、相关工作

1. 目前已有的深度学习配准方法，绝大多数使用了人工生成的 warp-field 作为配准的 ground-truth，即监督学习。
2. 有两篇文章([9]、[10])使用了非监督方法，然而均停留在较小尺度（二维切片/三维小区域）。而本文直接输入整个 CT 图像

四、模型

文章中使用了类似三维 U-net 的结构，将配准所用的两期图像直接输入网络，预测一个变换的向量场。注意，以往的配准大多是使用监督学习进行训练，即有变换向量场的 Ground Truth 样本；但本文是一种非监督的方法，没有这些样本，仅仅使用了原始数据。

主网络是类似三维 U-net 结构。U-net 为 U 型全卷积神经网络，目前是医学影像应用中广泛使用的分割网络。

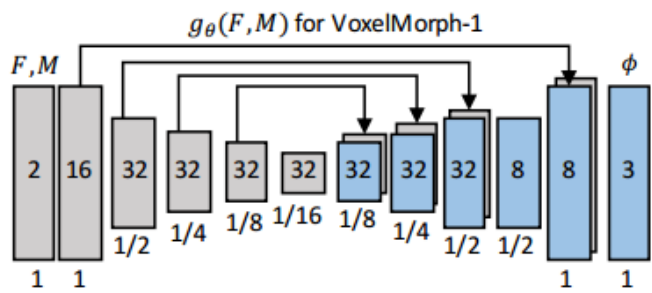


图 2 主干网络

使用三维的全卷积网络。特征如下：

- 1) 分为收缩与扩张两个路径，通道数不断减半/加倍。
- 2) 收缩与扩张路径中间有 concat（主网络结构图上方的箭头，将不同层的特征叠加）。这样可以同时提取不同尺度的信息。

工作流程如下：

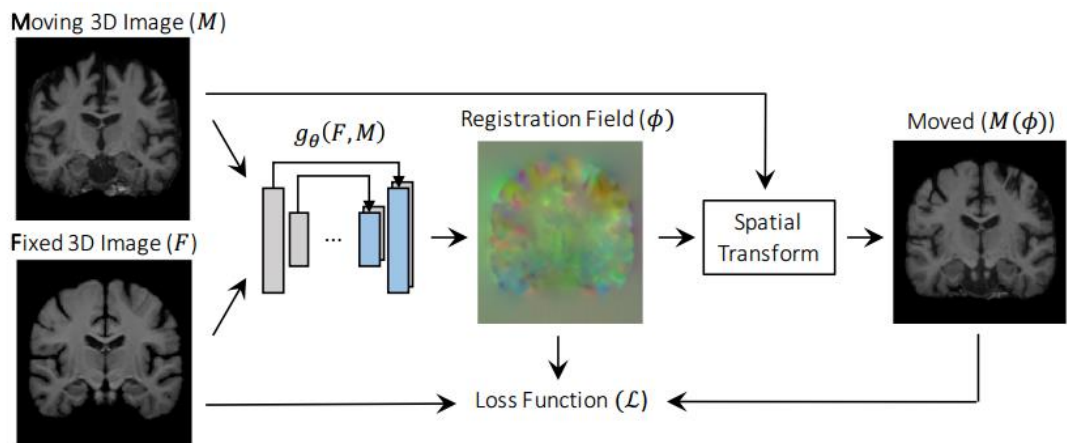


Figure 2: Overview of our method. We learn parameters for a function g that registers one 3D volume (M) to a second, fixed volume (F). During training, we warp M with ϕ using a spatial transformer function. Our loss compares M_ϕ and F and enforces smoothness of ϕ .

图 3 工作流程。将固定图像与变化图像输入三维 U-net，U-net 输出为两图之间的变换场 (Registration field，即点之间对应关系)。将变换场与变化图像结合，成为配准图像。配准图像应与固定图像越相似越好。

本文的损失函数如下：

$$\hat{\phi} = \arg \min_{\phi} \mathcal{L}(F, M, \phi), \quad (1)$$

where

$$\mathcal{L}(F, M, \phi) = \mathcal{L}_{sim}(F, M(\phi)) + \lambda \mathcal{L}_{smooth}(\phi), \quad (2)$$

损失函数包含两部分：第一部分是固定图像 F 和配准后图像 $M(\phi)$ 之间的差距（用相关系数等方法进行度量）第二部分是变化向量场 ϕ 的梯度，保证变化向量不会过于剧烈。这是一个正则项。总的来说，是传统的相似度量函数+光滑度量函数的联合学习。

$$\mathcal{L}(F, M, \phi) = -CC(F, M(\phi)) + \lambda \sum_{p \in \Omega} \|\nabla \phi(p)\|^2, \quad (7)$$

其中，相似度量函数用的是 Cross Correlation。

五、作者的测试结果

作者采用了大量数据集结合进行训练：ADNI [1], OASIS [2], ABIDE [3], ADHD200 [4], MCIC [5], PPMI [6], HABS [7], Harvard GSP [8], 以上均为脑部 MRI 数据集。

作者将不同数据集中的 MRI 数据均使用软件 FreeSurfer 进行预处理，分割脑部；并切分为同样大小（160*192*224）。

数据集大小：7329 train, 250 validation, 250 test （在医学图像人工智能中，5000+ 的数据可以说是非常难得的了。）

作者使用本文方法与专业配准软件 ANTs(Advanced normalization tools) 进行对比。ANTs 使用层次性的方法进行配准（Affine->Bspline，多次迭代）。测试结果如下：

Table 1: Average Dice scores and runtime results for affine alignment, ANTs, VoxelMorph-1, VoxelMorph-2. Standard deviations are in parentheses. The average Dice score is computed over all structures and subjects. Timing is computed after preprocessing. Our networks yield comparable results to ANTs in Dice score, while operating orders of magnitude faster during testing. To our knowledge, ANTs does not have a GPU implementation.

Method	Avg. Dice	GPU sec	CPU sec
Affine only	0.567 (0.157)	0	0
ANTs	0.749 (0.135)	-	9059 (2023)
VoxelMorph-1	0.742 (0.139)	0.365 (0.012)	57(1)
VoxelMorph-2	0.750 (0.137)	0.554 (0.017)	144 (1)

评价：可以看出，本文的配准结果与专业配准软件（代表传统方法）相比相差不多，但是速度有显著优势。本文可能有一个弱点，就是需要大量数据（作者找寻了近 20 个数据集进行训练）。但是，与监督学习方法相比，大量未处理数据的获取要容易的多（由于医生人工进行标注的时间成本非常昂贵）。

六、自己的实现（在肝脏数据集上测试）

更换脑部 MRI 数据为肝脏 CT 数据。数据来源：麻省总医院（来自于导师）
（训练集大小：74 个 case 的三期肝脏 CT 扫描。）

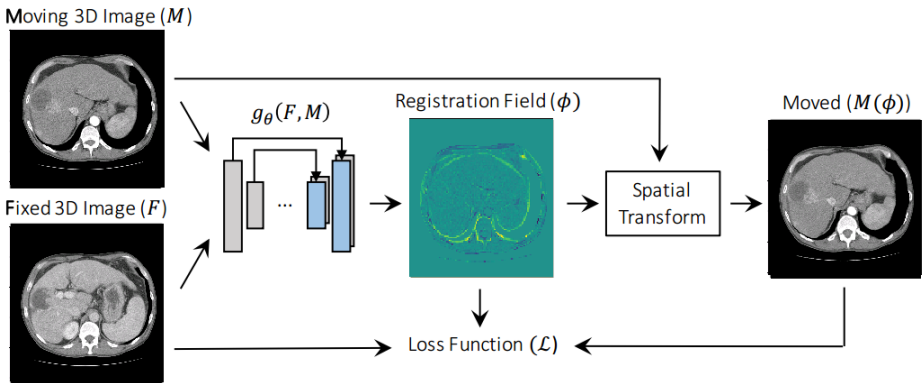
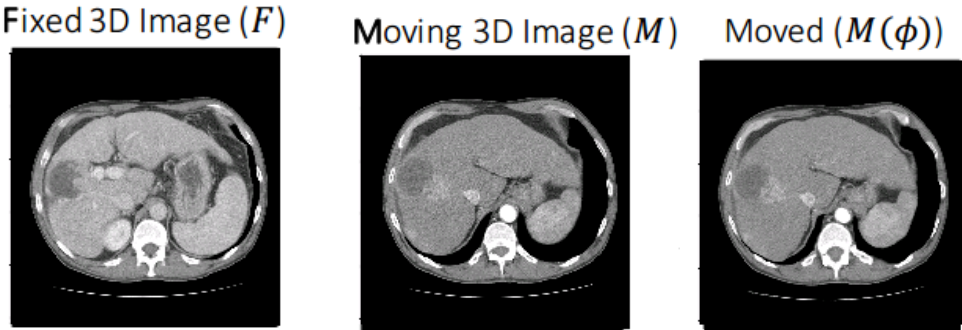


图 4 肝脏配准网络（图已经更换）

以图像的互信息（使用上文提到的标准配准工具 ANTs 计算，它由 University of Pennsylvania 开发）作为度量。以下面的 case 为例：



配准前互信息 $I_{FM} = 0.75$ ，配准后互信息 $I_{F(M\phi)} = 0.94$ （相似度增强）。（可从高亮脊椎骨区域观察配准前后变化）

七、讨论

测试后有以下发现：

在分辨率较低（如，将原有 512*512*512 的图像采样为 64*64*64 后）时，该网络经测试，可以较快的收敛，配准效果很好。

在分辨率较高时，训练速度明显变慢，配准输出图与标准图的相似性增加也较慢。可能是由于特征过多造成的。

如果有更多数据，有望对结果有较大提升。

测试发现配准结果对正则化参数 λ 的敏感性强。（即，图像相似性和变换向量场梯度在 Loss 中所占比重）对于不同的数据集（即边缘明显程度不同的数据集），应该需要不断的用试错法寻找最佳 Loss。

参考文献:

- [1] Susanne G Mueller, Michael W Weiner, Leon J Thal, Ronald C Petersen, Clifford R Jack, William Jagust, John Q Trojanowski, Arthur W Toga, and Laurel Beckett. Ways toward an early diagnosis in Alzheimers disease: the Alzheimers Disease Neuroimaging Initiative (ADNI). *Alzheimer's & Dementia*, 1(1):55–66, 2005.
- [2] Daniel S Marcus, Tracy H Wang, Jamie Parker, John G Csernansky, John C Morris, and Randy L Buckner. Open access series of imaging studies (oasis): cross-sectional mri data in young, middle aged, nondemented, and demented older adults. *Journal of cognitive neuroscience*, 19(9):1498–1507, 2007.
- [3] A. Di Martino et al. The autism brain imaging data exchange: towards a large-scale evaluation of the intrinsic brain architecture in autism. *Molecular psychiatry*, 19(6):659–667, 2014.
- [4] Michael P Milham, Damien Fair, Maarten Mennes, Stewart HMD Mostofsky, et al. The ADHD-200 consortium: a model to advance the translational potential of neuroimaging in clinical neuroscience. *Frontiers in systems neuroscience*, 6:62, 2012.
- [5] Randy L Gollub, Jody M Shoemaker, Margaret D King, Tonya White, Stefan Ehrlich, Scott R Sponheim, Vincent P Clark, Jessica A Turner, Bryon A Mueller, Vince Magnotta, et al. The MCIC collection: a shared repository of multi-modal, multi-site brain image data from a clinical investigation of schizophrenia. *Neuroinformatics*, 11(3):367–388, 2013.
- [6] Kenneth Marek, Danna Jennings, Shirley Lasch, Andrew Siderowf, Caroline Tanner, Tanya Simuni, Chris Coffey, Karl Kieburtz, Emily Flagg, Sohini Chowdhury, et al. The parkinson progression marker initiative (ppmi). *Progress in neurobiology*, 95(4):629–635, 2011.
- [7] Alexander Dagley, Molly LaPoint, Willem Huijbers, Trey Hedden, Donald G McLaren, Jasmeer P Chatwal, Kathryn V Papp, Rebecca E Amariglio, Deborah Blacker, Dorene M Rentz, et al. Harvard aging brain study: dataset and accessibility. *NeuroImage*, 2015.
- [8] Avram J Holmes, Marisa O Hollinshead, Timothy M OKeefe, Victor I Petrov, Gabriele R Fariello, Lawrence L Wald, Bruce Fischl, Bruce R Rosen, Ross W Mair, Joshua L Roffman, et al. Brain genomics superstruct project initial data release with structural, functional, and behavioral measures. *Scientific data*, 2, 2015.
- [9] Bob D de Vos, Floris F Berendsen, Max A Viergever, Marius Staring, and Ivana Išgum. End-to-end unsupervised deformable image registration with a convolutional neural network. In *DLMIA*, pages 204–212. Springer, 2017.
- [10] H. Li and Y. Fan. Non-rigid image registration using fully convolutional networks with deep self-supervision. *arXiv preprint arXiv:1709.00799*, 2017.