

# VoxelMorph based Normalization in the Prediction of Stable VS Progressive MCI Conversion with Convolutional Neural Networks

Daniel Sierra-Tome, Ubaldo Ramon-Julvez, Mónica Hernández, Elvira Mayordomo

Afiliación: *Computer Science for Complex System Modeling (COS2MOS)*  
Instituto de Investigación en Ingeniería de Aragón (I3A)  
Universidad de Zaragoza, Mariano Esquillor s/n, 50018, Zaragoza, Spain.  
Tel. +34-976762707, e-mail: [d.sierra@unizar.es](mailto:d.sierra@unizar.es)

## Resumen

El presente trabajo tiene como objetivo estudiar el efecto de la normalización en el sistema de clasificación de Spasov2019 [2] para el problema de deterioro cognitivo leve estable vs progresivo (sMCI vs pMCI), cuando esta normalización proviene de un algoritmo de deep-learning. Se ha elegido VoxelMorph [8], por ser uno de los algoritmos más utilizados como benchmark desde la introducción del deep-learning en el registro deformable de imágenes médicas. Se realizará una comparativa entre los métodos de registro difeomorfo tradicionales y VoxelMorph para este problema.

## Cuerpo

### Registro de imágenes médicas. Difeomorfismos. Normalización espacial.

El registro de imágenes médicas es uno de los problemas fundamentales en el análisis de imágenes médicas que permite establecer una correspondencia densa entre parejas de imágenes [1]. Una de las aplicaciones del registro es la normalización espacial a un sistema de referencia común (atlas) de forma que las estructuras anatómicas estén suficientemente alineadas para que puedan ser utilizadas como variables de entrada en sistemas de clasificación y modelos predictivos. Recientemente se ha demostrado empíricamente que esta normalización puede ser beneficiosa incrementando la precisión, sensibilidad y especificidad de sistemas de inteligencia artificial para establecer el diagnóstico y pronóstico de la enfermedad de Alzheimer [2-4].

El proceso de normalización espacial mediante registro consta de dos pasos: en primer lugar se realizan una serie de transformaciones afines, para evitar problemas de desalineamientos severos, debidos a los parámetros de adquisición del escáner

y la posición del paciente respecto al sistema de referencia de la máquina. En segundo lugar, se realiza una transformación no-rígida de muchos más grados de libertad. Los métodos de registro mediante difeomorfismos (e.g. transformaciones suaves con inversa suave) son preferibles, pues su modelo de deformación es coherente con el modelo físico de deformación que existiría entre las imágenes si se considerasen impresas en un medio deformable.

Los difeomorfismos poseen la propiedad de que el determinante de la matriz Jacobiana (matriz de derivadas parciales de la transformación respecto a las coordenadas cartesianas) se mantiene siempre positivo, indicando que el dominio de la imagen transformada no se dobla sobre sí mismo (*folding*). Como consecuencia, la topología de las estructuras anatómicas presentes en la imagen deformada se conserva respecto a la topología original.

Los métodos de registro difeomorfo tradicionales plantean el problema de registro a partir de la minimización de un funcional de energía que mide la similitud de las imágenes después del registro y la suavidad de la transformación difeomorfa estimada. Este problema se resuelve mediante un proceso de optimización basado en gradiente descendente, Gauss-Newton, u otros métodos tradicionales de optimización [5, 6].

Una de las principales limitaciones de los métodos de registro difeomorfo tradicionales es la gran carga computacional en tiempo y memoria [6]. Aunque se han propuesto diferentes combinaciones metodológicas muy interesantes, como la parametrización estacionaria o la parametrización en el espacio de funciones de banda limitada [6, 7], que unidas a la realización de los cálculos en la GPU han conseguido reducir la carga computacional, los métodos tradicionales vienen siendo desbancados por los métodos basados en deep-learning que

consiguen realizar inferencias en tiempos del orden de 1 segundo en volúmenes 3D.

## VoxelMorph

Este trabajo se centra en el estudio de las prestaciones de VoxelMorph [8] como algoritmo de normalización espacial en el problema de deterioro cognitivo leve estable vs progresivo (sMCI vs pMCI) con la arquitectura de Spasov2019 [2]. VoxelMorph es una suite de métodos de registro no rígido cuyo código abierto se encuentra disponible en <https://github.com/voxelmorph/voxelmorph>. Existen varias versiones de VoxelMorph y dos de ellas son difeomorfos (en teoría, empíricamente se ha demostrado la propensión de VoxelMorph a presentar *foldings*). El registro no rígido se plantea mediante la minimización de un funcional de energía similar al del problema LDDMM donde el mínimo se parametriza mediante una red neuronal convolucional. Dicha red convolucional está diseñada con una arquitectura encoder-decoder, similar a la presentada en UNet. [12].

Así, VoxelMorph se formula de una forma similar a un método de registro tradicional aunque la optimización se resuelve sobre los parámetros de la red neuronal mediante un gradiente descendente estocástico sobre el conjunto de entrenamiento utilizado (*data-driven approach*). Este es uno de los primeros métodos no supervisados de deep-learning propuestos para resolver el problema de registro.

Una vez que los problemas inherentes al entrenamiento de una red en un problema de estas características han sido resueltos, la red representa una función que permite inferir la deformación existente entre nuevas parejas de imágenes de acuerdo al modelo estimado durante el entrenamiento. Si las características del par de imágenes son similares a las del conjunto de entrenamiento, VoxelMorph puede ofrecer unos resultados con una precisión en la evaluación sobre segmentaciones manuales comparable al estado del arte con un tiempo de ejecución sobre la GPU de alrededor de 1 segundo.

## Problema sMCI vs pMCI.

### Sistema Spasov2019.

El problema sMCI vs pMCI consiste en identificar qué pacientes con deterioro cognitivo leve (MCI) desarrollarán la enfermedad de Alzheimer en un cierto periodo de tiempo. La solución de este problema es más compleja que el problema AD vs

CN (enfermos vs controles) y es de especial relevancia en la clínica, pues permite identificar pacientes susceptibles de recibir tratamientos experimentales.

En la literatura se han propuesto diferentes sistemas de inteligencia artificial para obtener sistemas de ayuda al diagnóstico y pronóstico basados en la solución de este problema [4]. Los más prometedores hasta la fecha son los métodos basados en deep-learning.

Entre ellos, destaca la arquitectura de Spasov2019, que propone la minimización simultánea del error de las dos tareas de clasificación: AD vs CN y sMCI vs pMCI [2]. Esta estrategia, denominada *joint learning*, tiene el objetivo de incrementar el conjunto de entrenamiento sobre el que estimar los pesos en las primeras capas de la red y de esta forma luchar contra el *overfitting*. La red combina datos clínicos, cognitivos, y de imágenes por resonancia magnética (MRI) normalizadas espacialmente.

## Contribución

En este trabajo se ha estudiado el efecto de la normalización espacial en el sistema de clasificación de Spasov2019 para el problema sMCI vs pMCI, cuando esta normalización proviene de VoxelMorph. Se ha realizado inferencia sobre un modelo entrenado sobre la base de datos de OASIS proporcionada por el Learn2Reg challenge (<https://learn2reg.grand-challenge.org/>), para evitar problemas potenciales de *data leakage*, pues los modelos públicos de VoxelMorph incluyen registros sobre ADNI.

## Resultados

### Datos utilizados

Los conjunto de datos de estudio fueron obtenidos de la base de datos de Alzheimer's Disease Neuroimaging Initiative (ADNI, <http://adni.loni.usc.edu>), organización cuya motivación y objetivo es proporcionar imágenes médicas, biomarcadores e información clínica que permitan ayudar hacia un mayor conocimiento de la evolución del deterioro cognitivo y las etapas tempranas de la enfermedad de Alzheimer. Con el fin de mantener la reproducibilidad y comparación de los resultados, los autores de [2] nos proporcionaron los individuos exactos utilizados, formados por 435 hombres y 350 mujeres con edades comprendidas entre los 55 y 91 años. En este conjunto de datos, 191 sujetos pertenecen a AD, 181 a HC, 179 a pMCI y 227 a sMCI.

Las imágenes MRI-T1 originales utilizadas en el estudio fueron las correspondientes a la visita inicial (baseline) de cada uno de los individuos, preprocesadas según un conjunto estándar de operaciones para estudios sobre imágenes neurológicas propuesto en [9]. Las imágenes fueron reorientadas y normalizadas afinmente al atlas MNI152 mediante la librería FSL (<https://fsl.fmrib.ox.ac.uk>). Posteriormente, la distorsión de las imágenes MRI inducida por el campo magnético fue eliminada mediante el algoritmo de ANTS N3 *bias field correction* (<http://stnava.github.io/ANTs>) y el cráneo fue suprimido con la herramienta Robbex (<https://www.nitrc.org/projects/robbex>). Por último se realizó la normalización espacial al atlas MNI152 con los métodos de registro tradicionales considerados como baseline en este estudio y con Voxelmorph.

Los métodos de registro tradicionales considerados en el estudio son registro afin, ANTS [10], y BL PDE-LDDMM [11].

## Resultados

La Figura 1a muestra los diagramas de caja de la precisión de acierto en la tarea de clasificación sMCI vs pMCI para un total de 10 experimentos utilizando únicamente las imágenes normalizadas como datos de entrada. La Figura 1b muestra los diagramas de caja de la precisión cuando las imágenes normalizadas se combinan con la información clínica utilizada en el sistema de Spasov19.

Los modelos entrenados con MRIs muestran unos resultados ligeramente superiores para los casos de uso de las imágenes registradas, tanto para el método tradicional BL PDE-LDDMM NCC como para Voxelmorph. Las diferencias no son significativas entre ambos métodos de registro, mostrando la capacidad de VoxelMorph de obtener una precisión a la altura de los mejores métodos tradicionales.

Los modelos entrenados con información clínica e imágenes obtienen en general una mejora significativa en su capacidad predictiva, debido a que la información clínica contiene variables bastante explicativas del diagnóstico de los pacientes. Es sorprendente cómo la normalización afin alcanza una precisión comparable a la del mejor método tradicional. En este caso VoxelMorph presenta unos resultados ligeramente inferiores.

## Conclusiones

En este trabajo se ha estudiado el efecto de la normalización espacial de VoxelMorph en el sistema de clasificación de Spasov2019 para el problema sMCI vs pMCI. Los resultados de la precisión del método son comparables al mejor método de normalización espacial basado en métodos tradicionales. Aunque preliminar, este estudio arroja unas conclusiones y unas direcciones de investigación muy interesantes para ser realizadas en el futuro.

En primer lugar, se ha mostrado cómo un modelo basado en aprendizaje no supervisado es capaz de obtener una precisión similar a la de los mejores métodos tradicionales. Este resultado junto con la eficiencia en tiempo puede convertirlo en el algoritmo de elección para el diseño de sistemas de ayuda al diagnóstico. En segundo lugar, la combinación de datos clínicos con imagen parecen mejorar la precisión de los sistemas pero difumina el efecto de la normalización espacial en la capacidad predictiva. Este efecto será estudiado como trabajo futuro utilizando técnicas de interpretabilidad.

## Agradecimientos

Este trabajo ha sido parcialmente financiado por los proyectos nacionales TIN2016-80347-R (DIAMOND) y PID2019-104358RB-I00 (DL-Ageing), y el grupo de referencia del Gobierno de Aragón T 64 20R (COS2MOS).

## REFERENCIAS

- [1]. A. Sotiras, C. Davatzikos y N. Paragios, "Deformable Medical Image Registration: A Survey," in *IEEE Transactions on Medical Imaging*, vol. 32, no. 7, pp. 1153-1190, July 2013, doi: 10.1109/TMI.2013.2265603.
- [2]. S. E. Spasov, L. Passamonti, A. Duggento, P. Lio, N. Toschi, and ADNI, "A parameter-efficient deep learning approach to predict conversion from mild cognitive impairment to Alzheimer's disease", *Neuroimage*, vol. 189, pp. 276-287, 2019.
- [3]. U. Ramon-Julvez, M. Hernandez, E. Mayordomo y Adni, "Analysis of the Influence of Diffeomorphic Normalization in the Prediction of Stable VS Progressive MCI Conversion with Convolutional Neural Networks," *2020 IEEE 17th International Symposium on Biomedical Imaging (ISBI)*, 2020, pp. 1120-1124, doi: 10.1109/ISBI45749.2020.9098445.
- [4]. J. Wen, E. Thibeau-Sutre, M. Diaz-Melo, J. Samper-Gonzalez, A. Routier, S. Bottani, et al, "Convolutional neural networks for classification of Alzheimer's Disease: overview and reproducible evaluation,

Medical Image Analysis, vol. 63, 2020, doi: 10.1016/j.media.2020.101694

- [5]. M. F. Beg, M.I. Miller, A. Trounev, L. Younes, “Computing large deformation metric mappings via geodesic flows of diffeomorphisms” International Journal of Computer Vision, vol. 61, 2005, pp. 139-157, doi: 10.1023/B:VISI.0000043755.93987.a
- [6]. M. Hernandez. “A comparative study of different variants of Newton–Krylov PDE-Constrained Stokes-LDDMM Parametrized in the Space of Band-Limited Vector Fields, SIAM Journal of Imaging Science, vol. 12, 2019, doi: 10.1137/18M1195310
- [7]. M. Zhang, P.T. Fletcher, “Fast diffeomorphic Image Registration via Fourier-Approximated Lie Algebras”, International Journal of Computer Vision, vol. 127, pp. 61-73, 2019, doi:10.1007/s11263-018-1099-x
- [8]. G. Balakrishnan, A. Zhao, M.R. Sabuncu, J. Guttag y A.V. Dalca. “VoxelMorph: A learning framework for deformable medical image registration”. IEEE Transactions on Medical Imaging. vol. 38, pp. 1788-1800, 2019, doi:10.1109/tmi.2019.2897538

- [9]. M. HadjHamou, M. Lorenzi, N. Ayache and X. Pennec, “Longitudinal analysis of image time series with diffeomorphic deformations: a computational framework based on stationary velocity fields”, Frontiers in Neuroscience, vol 10, 2016. doi: 10.3389/fnins.2016.00236
- [10]. B.B. Avants, C.L Epstein, M. Grossman y J.C. Gee, “Symmetric diffeomorphic image registration with cross-correlation: evaluating automated labeling of elderly and neurodegenerative brain. Medical Analysis, vol. 12, pp. 26-41, 2008, doi: 10.1016/j.media.2007.06.004
- [11]. M. Hernandez, U. Ramon-Julvez, D. Sierra Tome, “Partial differential equation-constrained diffeomorphic registration from sum of squared differences to normalized cross-correlation, normalized gradient fields, and mutual information: a unifying framework. Sensors, vol. 22, 2022, doi: 10.3390/s22103735
- [12]. P. Isola, J. Zhu, T. Zhou, A.A. Efros, “Image-to-Image translation with conditional adversarial networks”, 2016, doi: 10.48550/ARXIV.1611.07004

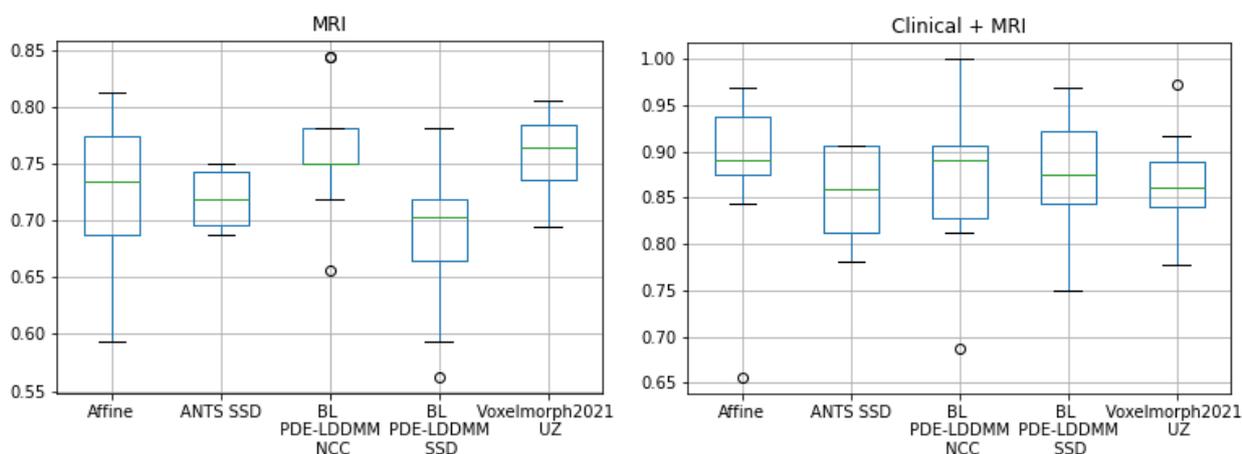


Figura 1: Diagramas de caja de los valores de precisión (*accuracy*) obtenidos por cada método de normalización espacial utilizado. La Figura 1a (izquierda) muestra los valores del sistema entrenado únicamente con imágenes MRI como datos de entrada. La Figura 1b (derecha) muestra los valores del sistema entrenado con imágenes e información clínica.