Critical Review: B. B. Avants, C. L. Epstein, M. Grossman, and J. C. Gee, "**Symmetric Diffeomorphic Image Registration with Cross-Correlation:** Evaluating Automated Labeling of Elderly and Neurodegerative Brain," Medical Image Analysis, Vol. 12, No. 1, pp.26-41, 2008.

Avants et al.'s goal in writing this paper is to propose a new deformable registration method and compare it to existing methods using brain MRI data. The method they propose is called symmetric image normalization (SyN). The method is meant to achieve better registration by maximizing cross correlation within the space of diffeomorphic maps, and the authors provide the Euler-Lagrange equations necessary to achieve this. SyN is advantageous in that it guarantees identical results each time the same two images are registered, and it takes advantage of exact inverse transformations guaranteed by diffeomorphisms. This method is most unique by the fact that cross correlation has not been investigated in diffeomorphic registrations. Such a combination allows for the possibility of symmetrizing cross correlation Euler-Lagrange equations. The authors test their method against the elastic method and the ITK implementation of Thirion's Demons algorithm.

The paper then goes on to summarize the details of the Demons method, the SyN method, and the SyN method with cross correlation. The Demons method is different in that rather than using a function to update the image, it uses optical flow. More specifically, it uses an elastic regularizer to solve an optical flow problem. Elastic regularization is a regression method using regularization to combine penalties of lasso and ridge least squares methods. Here, one image is fixed and the other moves by bringing its level sets into correspondence with the fixed image. The optical flow term is iteratively computed and added to total displacement (init. 0) which is then smoothed with Gaussian filter. It has been shown that there is agreement between Demons and manual labeling of image structures using the Dice statistic.

SyN is a symmetric diffeomorphism; it is constrained by the authors to diffeomorphic space $Diff_0$ with homogenous boundary conditions meaning the image border maps to itself. Most methods are symmetric in theory but not implementation. The use of diffeomorphism allows genuine symmetry and the ability to implement as such. Because it is symmetric, it is guaranteed that the path from Image I to Image J is the same as when computed from J to I for all similarity metrics. For other methods, symmetrization of the Jacobian is possible, but only when the similarity metric is intensity. Symmetry also eliminates the problem of results depending on which image is chosen to be fixed and which moves. Further, because it is invertible, it guarantees sub-pixel accuracy of invertible transformations in the discrete domain. What is also unique about SyN is that rather than having a fixed and moving image, there is a series of diffeomorphisms that connect two images, and both images deform in time along the series. Diffeomorphisms can be decomposed into two parts, which allowed the authors to define a variational energy that divided the diffeomorphisms in half. Thus, the path length and deformation is divided evenly between the two images, allowing correspondence to be found with equal consideration of both images. The full path and inverse is found by adding the two paths (see Fig. 1 below).

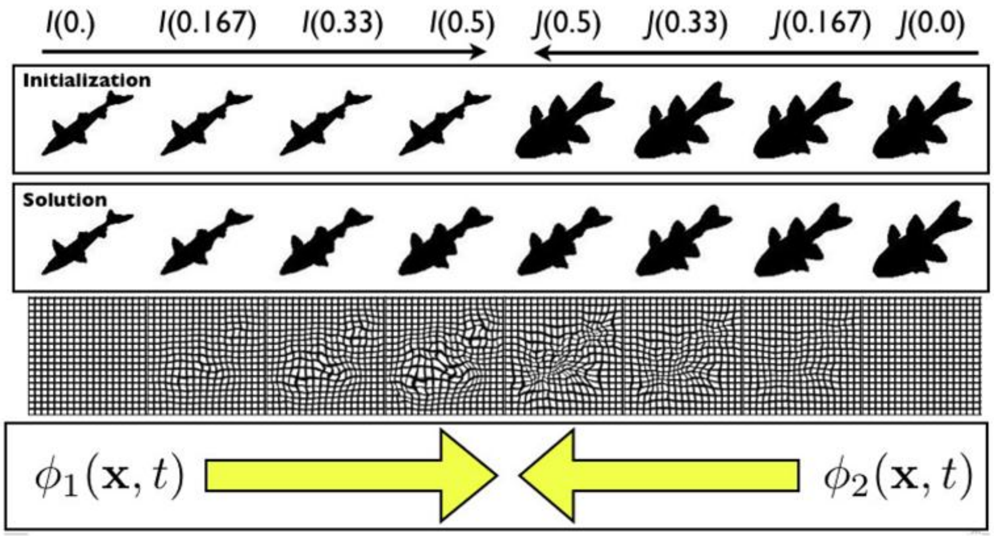


Figure 1. Source: Avants et al.

The method advocated by the authors is not just general SyN, but SyN where the similarity metric is cross correlation. The authors note here the difference between SyN with cross correlation and the elastic method, which also uses cross correlation: elastic method uses a balance constraint between regularization term and similarity term, whereas SyN with cross correlation uses unconstrained optimization of the similarity term within the space of diffeomorphisms. In any case, cross correlation is advantageous because it adapts to locally varying intensities, it depends only on estimates of local image average and variance (can be measured with few samples), and it is robust to unpredictable illumination, reflectance, etc. Mathematically, the cross correlation term is:

$$CC(\bar{I},\bar{J},\mathbf{x}) = \frac{<\bar{I},\bar{J}>^2}{<\bar{I}><\bar{J}>} = A^2/BC,$$

and the optimization problem is:

$$E_{CC}(\bar{I},\bar{J}) = \inf_{\phi_1} \inf_{\phi_2} \int_{t=0}^{\frac{1}{2}} \left\{ \|\upsilon_1(\mathbf{x},t)\|_L^2 + \|\upsilon_2(\mathbf{x},t)\|_L^2 \right\} dt + \int_\Omega CC(\bar{I},\bar{J},\mathbf{x}) d\Omega.$$

$$\text{Subject to each } \phi_i \in Diff_0 \text{ the solution of: } d\phi_i(\mathbf{x},t)/dt = \upsilon_i(\phi_i(x,t),t) \text{ with } \phi_i(\mathbf{x},0) = \mathbf{Id} \text{ and } \phi_i^{-1}(\phi_i) = \mathbf{Id}, \phi_i(\phi_i^{-1}) = \mathbf{Id}.$$

Further, the Euler-Lagrange equations obtained by following Beg's derivation are:

$$\nabla_{\phi_1(\mathbf{x},0.5)} E_{CC}(\mathbf{x}), = 2L\upsilon_1(\mathbf{x},0.5) + \frac{2A}{BC}(\bar{J}(\mathbf{x}) - \frac{A}{B}\bar{I}(\mathbf{x}))\ |D\phi_1\ |\nabla\bar{I}(\mathbf{x}),$$

$$\nabla_{\phi_2(\mathbf{x},0.5)} E_{CC}(\mathbf{x}), = 2L\upsilon_2(\mathbf{x},0.5) + \frac{2A}{BC}(\bar{I}(\mathbf{x}) - \frac{A}{C}\bar{J}(\mathbf{x}))\ |D\phi_2\ |\nabla\bar{J}(\mathbf{x}).$$

The authors outline three algorithms for obtaining symmetric image normalization with cross correlation. Algorithm 1 allows rapid computation of the Euler-Lagrange equations simultaneously by finding and storing images A, B, and C (see CC definition above). Algorithm 2 is the LPF method, which is used to check that the spatiotemporal maps satisfy the ODE and invertibility constraints. This is done by pushing φ forward a small amount and performing gradient descent. Algorithm 3 is the overall algorithm for SyN with cross correlation. It iterates until convergence $\phi_1^{-1}(1)=\phi_2(1)=\phi_1^{-1}(\phi_2(\mathbf{x},0.5),0.5)$ using Algorithms 1 and 2 to compute cross correlation, the velocity field, and update the spatiotemporal maps.

In terms of implementation, the authors used the same ITK deformable registration framework as Demons, meaning both have the same code base. In terms of testing, the major difference between Demons and the elastic method is using cross correlation as the similarity metric, so we can see effect of using cross correlation simply by comparing these two method. SyN and the elastic method differ more fundamentally in their transformation models. Thus, we can evaluate the effectiveness of SyN by comparing it to the elastic method.

The data involved were 20 T1 MRI images, where 10 were elderly brains and 10 were FTD (frontotemporal dementia) brains. There was also a template brain with labels of the cortex, hippocampus, amygdala, and cerebellum. There were 60 deformable registrations (1 per image per method) between the test data and the template. Effectiveness is evaluated by Dice overlap ratios between automatic and manual (gold standard) structural segmentations. The ratio of running times was reported as Demons:elastic:SyN being 1:4.2:5.5.

The results are shown in the figures below. Figure 2 shows differences in details caught by SyN but not the other two methods. Figure 3 shows absolute value of image differences from the original, showing slightly better results for SyN. It also shows that some details are not captured by any of the methods. Figure 4 shows labeling and the errors between manual and automated labeling. Figure 5 shows the average Dice statistic for each structure. These results show that SyN outperforms the other two methods by consistently having a higher overlap ratio. Furthermore, a more exact comparison of volume measurements was obtained for three structures: the temporal, frontal, and parietal lobes. These results are shown below in Table 1, which lists Pearson correlations, and Table 2, which shows the absolute volume error, between manual and each method.

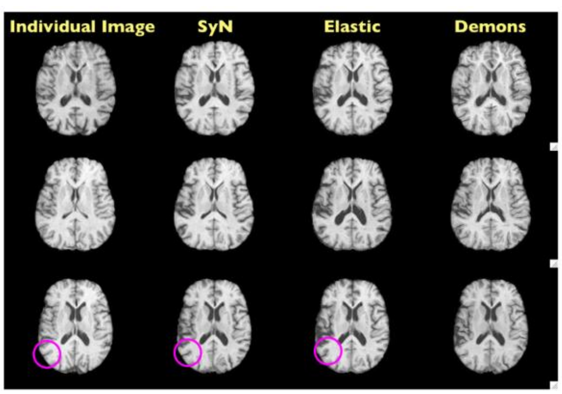


Figure 2. Avants et al.

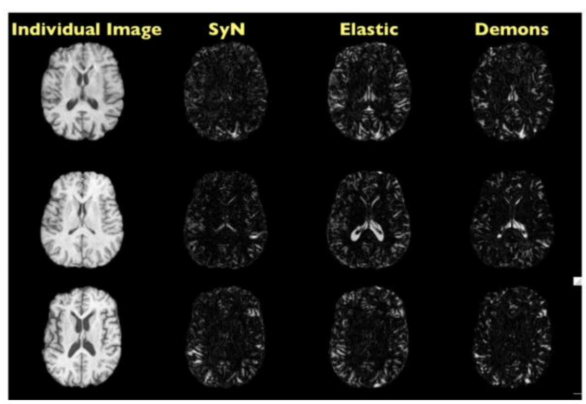


Figure 3. Avants et al.

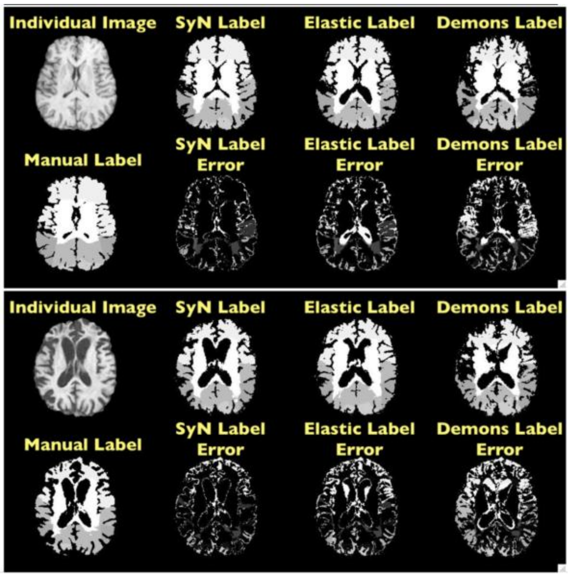
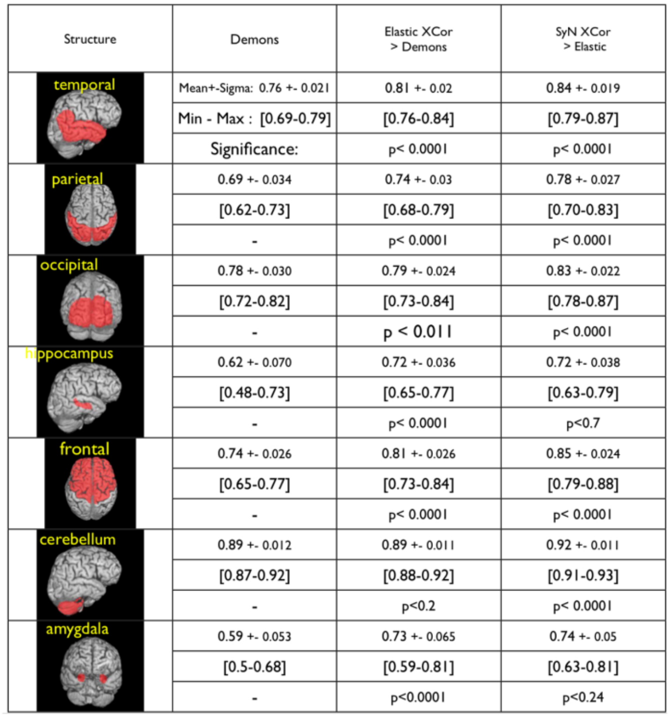


Figure 4. Avants et al.



Structure	Demons	Elastic XCor > Demons	SyN XCor > Elastic
temporal	Mean+-Sigma: 0.76 +- 0.021	0.81 +- 0.02	0.84 +- 0.019
	Min - Max : [0.69-0.79]	[0.76-0.84]	[0.79-0.87]
	Significance:	p< 0.0001	p< 0.0001
parietal	0.69 +- 0.034	0.74 +- 0.03	0.78 +- 0.027
	[0.62-0.73]	[0.68-0.79]	[0.70-0.83]
	-	p< 0.0001	p< 0.0001
occipital	0.78 +- 0.030	0.79 +- 0.024	0.83 +- 0.022
	[0.72-0.82]	[0.73-0.84]	[0.78-0.87]
	-	p < 0.011	p< 0.0001
hippocampus	0.62 +- 0.070	0.72 +- 0.036	0.72 +- 0.038
	[0.48-0.73]	[0.65-0.77]	[0.63-0.79]
	-	p< 0.0001	p<0.7
frontal	0.74 +- 0.026	0.81 +- 0.026	0.85 +- 0.024
	[0.65-0.77]	[0.73-0.84]	[0.79-0.88]
	-	p< 0.0001	p< 0.0001
cerebellum	0.89 +- 0.012	0.89 +- 0.011	0.92 +- 0.011
	[0.87-0.92]	[0.88-0.92]	[0.91-0.93]
	-	p<0.2	p< 0.0001
amygdala	0.59 +- 0.053	0.73 +- 0.065	0.74 +- 0.05
	[0.5-0.68]	[0.59-0.81]	[0.63-0.81]
	-	p<0.0001	p<0.24

Figure 5. Avants et al.

Table 1:

Structure	Corr(Man,Syn)	Corr(Man,Elas)	Corr(Man,Demon)
Temporal	0.86	0.69	0.79
Frontal	0.89	0.67	0.71
Parietal	0.71	0.42	0.66

Table 2:

Structure	VolErr(Man,Syn)	VolErr(Man,Elas)	VolErr(Man,Demons)
Temporal	8.4	9.2	8.7
Frontal	11.1	16.1	15.8
Parietal	7.9	9.3	7.9

There were several other findings. First, there is no significant difference between the minimum Jacobian of SyN vs the elastic method. This means that SyN's results are not significantly less constrained at a local level as was theorized. Second, there is no significant difference in volumes between FTD and elderly individuals. Third, automated methods tend to overestimate volumes. This is due to two things: segmentation bias toward template, and automated segmentations are smoothed while manual are not. Finally, though SyN outperforms other methods, it is still not able to claim accurate reproduction of manual labeling. This is because: 1) expert knowledge is not encoded in these methods, 2) uncertainty in neuroanatomical labeling limits accuracy in both automated and manual labeling, and 3) it is not known the extent to which different brains are diffeomorphic to each other.

I have several criticisms of this paper, both on its own as an academic work and in reference to its applicability to my project. First, the Dice statistic threshold is said to be set to 0.6 for small structures and 0.8 for large structures – must be above this to be considered “good” overlap. However, it seems these values were chosen arbitrarily. I would like there to be a more exact way of deciding what good overlap is. Second, the authors mention many times that their method is very fast, yet it is 5.5 times slower than the Demons algorithm. This is not a terrible drawback though because it is unlikely that physicians will need to immediately plan post-operative radiotherapy. Third, this method is good because for our project because it will work well with CT to CT registration, which is what we seek. Fourth, it may be more useful for us to use an algorithm where one image is fixed and the other moves. We are removing tissue, so we do not want a registered image that shows tissue that is halfway removed; it might be better to bias a fixed post-operative image with tissue removed and have the pre-operative image conform to it. Fifth, as mentioned before, there is typically an overestimation of volume by automated methods. This is undesirable, but it is still allowable. However, if volumes were underestimated, we could not use registration for this application since it is necessary that the whole volume containing remaining cancer cells be irradiated. In fact, a slight overestimation puts us on the safer side. Sixth, it was also mentioned before that SyN still cannot claim accurate volume representation. This challenges the utility of our project’s end goal since we are interested in localizing a volume. However, we are more interested in how a volume changes in general, rather than on specific values of volume. Regardless, this project will serve as a good first step as registration algorithms continue to improve.