AAPM 2016 JUL 31–AUG 4
COMMUNICATING OUR VALUE. IMPROVING OUR FUTURE.
58TH ANNUAL MEETING & EXHIBITION | WASHINGTON, DC

Parameterized Image Similarity for Fast, Automated DIR Accuracy Quantification

John Neylon, Yugang Min, Katelyn Hasse, Daniel Low, and Anand Santhanam
UCLA Radiation Oncology

UCLA Health Radiation Oncology

PURPOSE

Quantifying deformable image registration accuracy is a difficult task in a clinical setting due to poor image quality of the daily imaging modalities (CBCT, MVCT), and the lack of known ground truth deformations. The current gold standard is physical distance between manually placed landmarks, known as the target registration error. However, this process is time-consuming and subject to user-biases. Image similarity metrics (ISM) may provide an alternative way to represent the registration error, but need to be parameterized and translated to physical distance measures to **enable a fast, quantitative comparison of registration performance**.

MATERIALS & METHODS

- A biomechanical model was used to create pairs of volumetric images with known ground-truth deformation vector fields [1].
- A dense parameter space data set (2400 data points) was generated to fully characterize the relationship between the ISMs and the ground truth TRE (gt-TRE).
- A multi-pose anatomy training data set was generated to analyze performance across clinically realistic deformations. The biomechanical model was employed to induce 45 different postures, with 6 levels of tumor regression at each posture, which were then registered with 5 different sets of parameters (1350 data points per patient).
- A cost function was developed to test the relationship between the parameterized ISMs and the known registration error for sub-volumes enclosing critical radiotherapy structures in the head-and-neck region.
- Parameterization terms using ISMs were developed from two expectations of an accurate registration:
    - The warped data obtained from applying the deformation vector field (DVF) to the source data should closely match the target data
    - The similarity between the source and warped pair (I_{SW}) should match the similarity between the source and target pair (I_{ST})

$$CFR = fX^m - (1-f)Y^n \; ; \; X = 1 - |I_{ST} - I_{SW}| \; , \; Y = I_{TW}$$

- Similarity analysis was performed on each registration, examining several sub-volumes around regions of interest, including the parotids, cord, and PTV of head and neck patients.
- 25% of the ground truth data were also fed through a 3 layer neural network (8x11x4 neurons) to infer an estimate TRE (nn-TRE).
- Separate neural networks were trained for the dense parameter space data set and multi-pose anatomy data set.
- Multi-pose anatomy data sets were generated for 10 patients, and used to train a patient-specific neural network.

CONTACT INFORMATION: jneylon@mednet.ucla.edu

RESULTS

While the cost function (CFR) showed that a relationship could be established between the image similarity metrics and target registration error, the function was not sophisticated enough to fully characterize the relationship (Fig. 1,2).

The trained neural network (schematic representation shown in Fig. 5) provided the necessary level of abstraction, and achieved over 88% accuracy when trained on the dense parameter space data set (Fig. 3,4). Additionally, correlations of 0.9 or better were achieved for three of four contours investigated (Table 1), far outperforming the cost function.

For the multi-pose anatomy data set, over 95% accuracy was achieved by the neural network for the PTV sub-volume (Fig. 6), and over 90% for all structures. Table 2 summarizes 10 patient-specific neural network results trained on multi-pose anatomy data.

Table 1. Correlation with ground-truth TRE for cost function response and neural network predicted TRE

	CFR v. gt-TRE		nn-TRE v. gt-TRE
	Reference CFV	Best CFV	
PTV1	-0.467	-0.649	0.950
Left Parotid	-0.921	-0.952	0.988
Right Parotid	-0.860	-0.958	0.952
Cord	0.138	-0.109	0.753

Table 2. Summarized Results of neural network performance trained on multi-pose anatomy data sets generated from 10 patients

Pt	Accuracy		Correlation		
	%	mm	Lt Prtd	Rt Prtd	PTV
1	94.76	0.022	0.9452	0.9399	0.8924
2	82.43	0.249	0.9266	0.9214	0.9017
3	93.66	0.032	0.9654	0.9588	0.8986
4	94.28	0.022	0.9500	0.9571	0.8421
5	95.38	0.019	0.9862	0.9745	0.9740
6	95.25	0.021	0.9847	0.9504	0.9566
7	86.50	0.227	0.9560	0.9486	0.9526
8	86.91	0.134	0.9442	0.9401	0.9472
9	92.55	0.054	0.9568	0.9562	0.9015
10	91.61	0.057	0.9751	0.9722	0.7474
Avg	91.33	0.084	0.9590	0.9519	0.9014

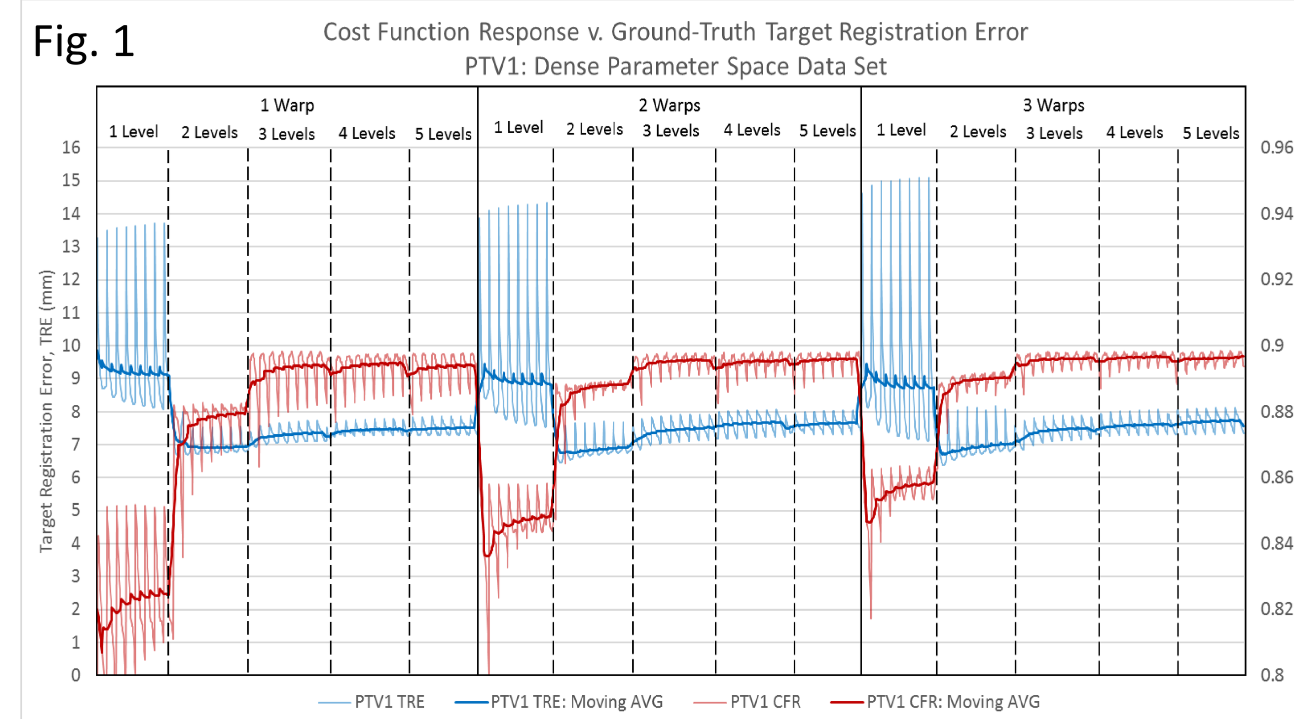
Fig. 1


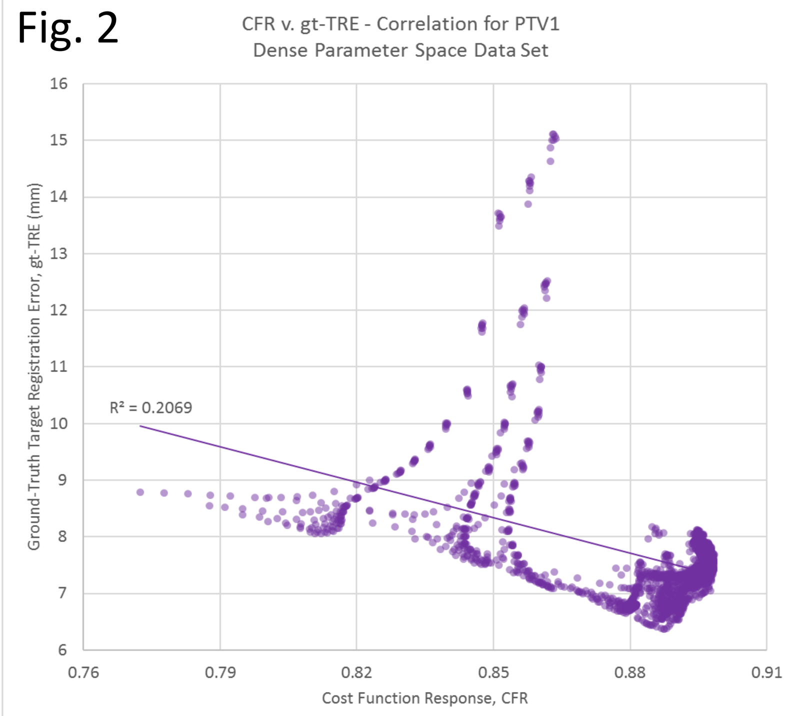
Fig. 2


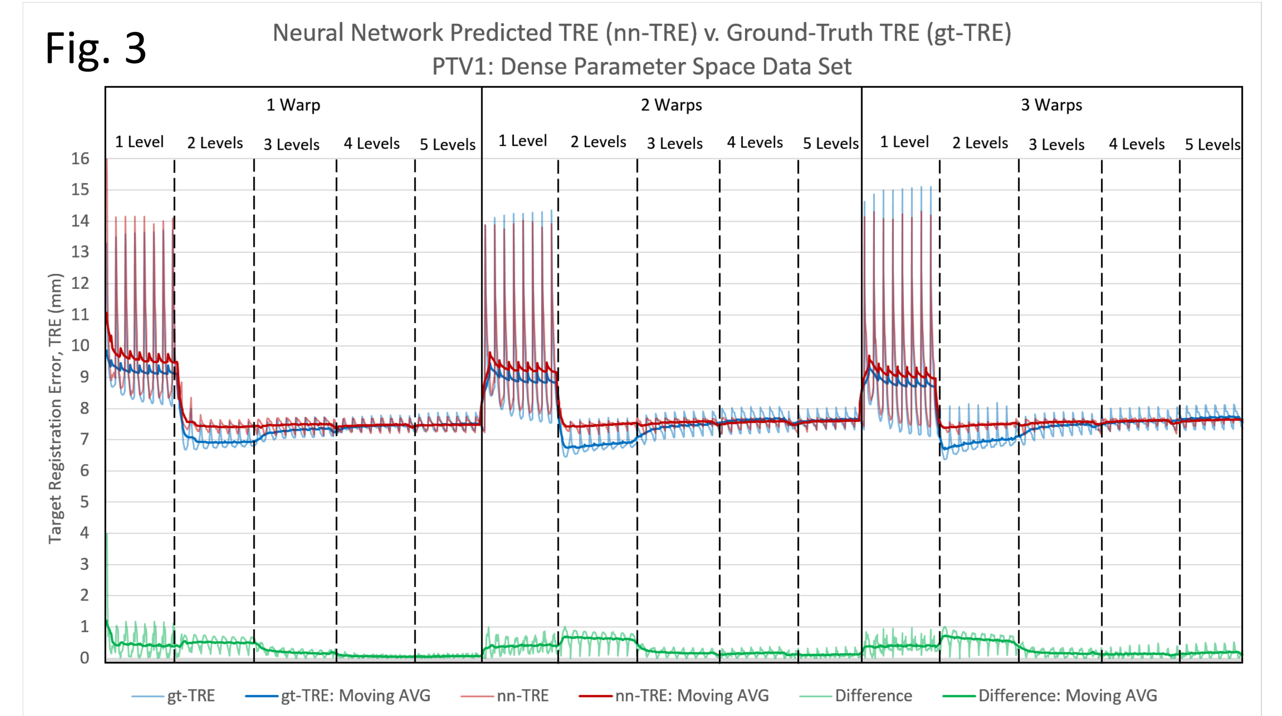
Fig. 3


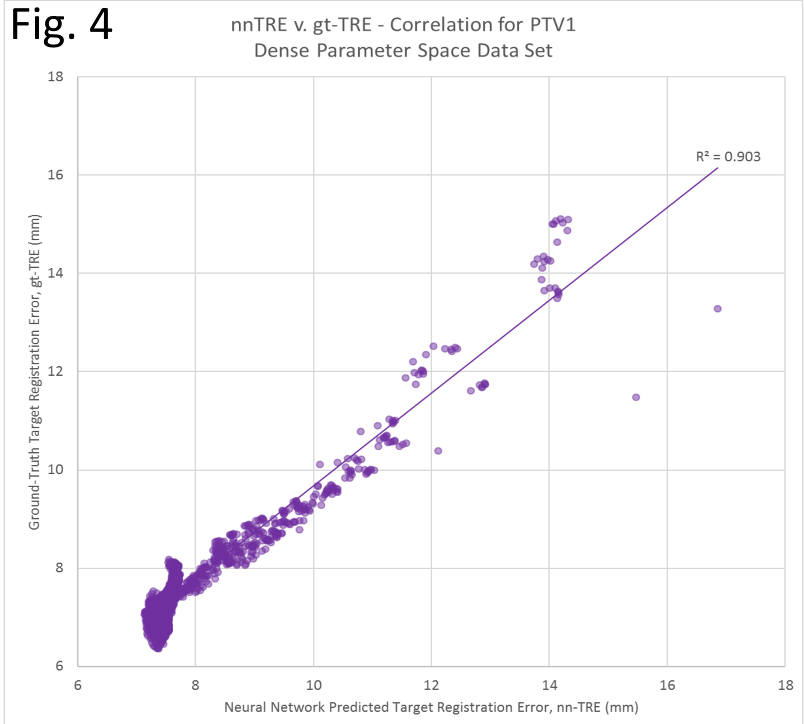
Fig. 4


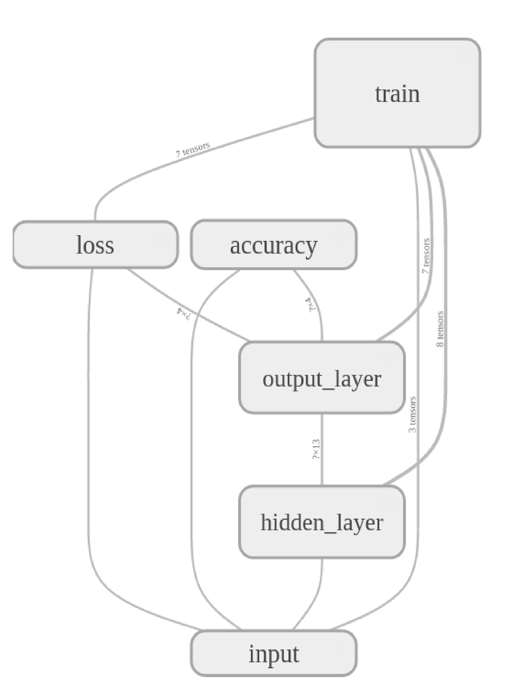
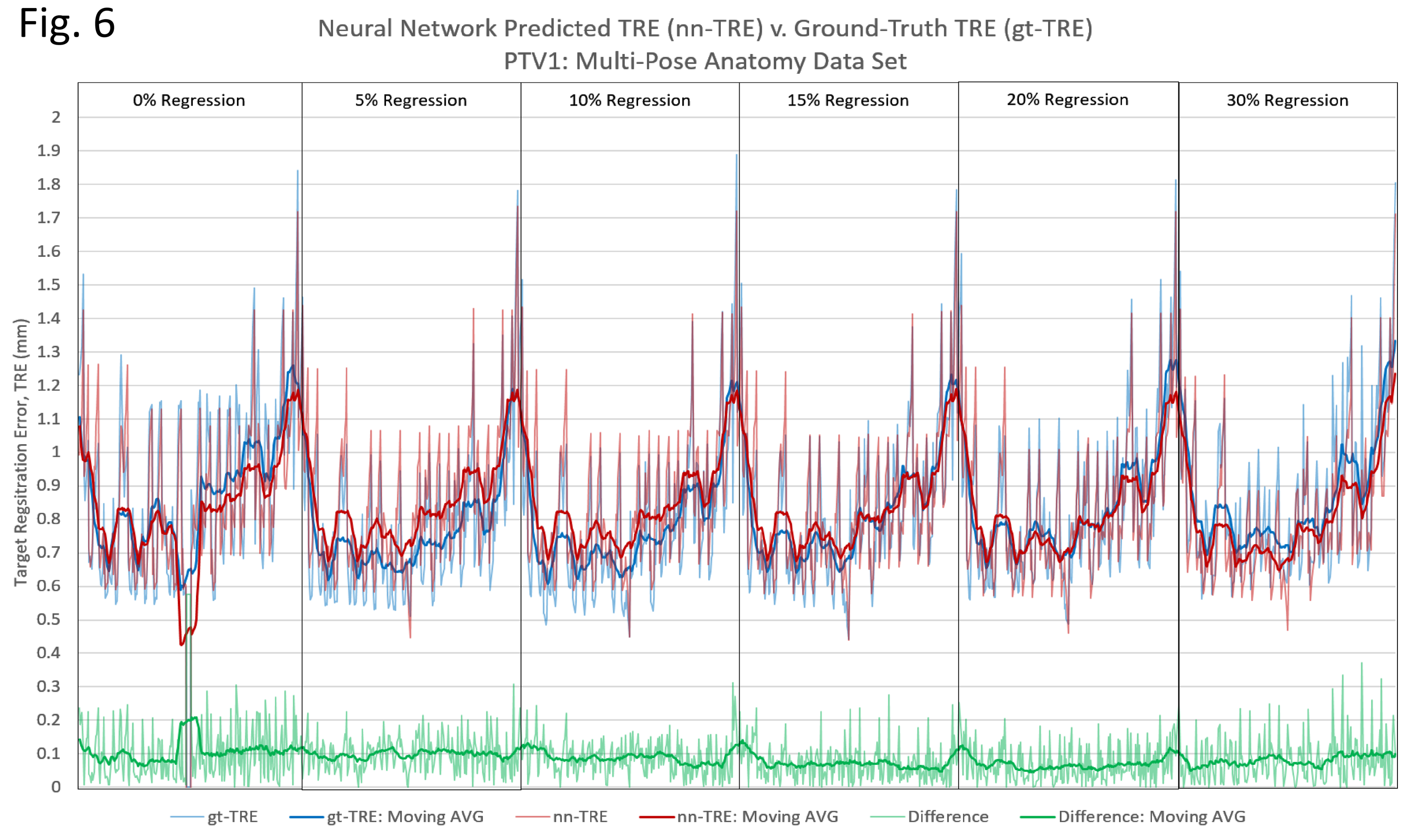
Fig. 5: NN Architecture


Fig. 6


CONCLUSIONS

The formulation presented demonstrates the ability for fast, accurate quantification of registration performance. **When sufficiently trained on annotated data, a neural network can learn to infer an estimate of target registration error from parameterized image similarity metrics**. Such networks have potential clinical impact in patient and site-specific optimization, and stream-lining clinical registration validation.

REFERENCES

[1] Neylon, J., Qi, X., Sheng, K., Staton, R., Pukala, J., Manon, R., Low, D. A., Kupelian, P., and Santhanam, A., "A GPU based high-resolution multilevel biomechanical head and neck model for validating deformable image registration," Med Phys 42(1), 232-43 (2015).