Title: A Shape Aware Loss for Semantic Segmentation

Authors: Francesco Calivà*, PhD, Claudia Iriondo*, Sharmila Majumdar, PhD, Valentina Pedoia, PhD.
Affiliations: Centre for Intelligent Imaging, Department of Radiology and Biomedical Imaging, University of California, San Francisco.

Introduction: Semantic segmentation of Magnetic Resonance Images (MRI) is crucial for the extraction of biomarkers indicative of the presence or severity of a disease. We focus our work on knee osteoarthritis, a degenerative joint disease affecting over 27 millions of Americans, of which the onset is manifested in the form of cartilage loss and changes in the subchondral and trabecular bone. Deep Learning methods applied to semantic segmentation require the tuning of trainable parameters; an operation generally driven by iteratively comparing the current segmentation against the desired one by means of an explicit loss function. Unfortunately, this approach can suffer from low performance at object boundaries. To overcome such limitations, in this work, we propose to use distance maps (ϕ) derived from ground truth masks (Figure 1-top), which guides the network's focus towards boundary regions that are difficult to segment.

Methods: We conducted an initial experiment on a subset of 40 DESS MRI from the Osteoarthritis Initiative (OAI) dataset, subdivided into training/validation/test set splits comprising 25/5/10 volumes respectively. In pursuance of segmenting patella, proximal tibia and distal femur, we trained an architecture resembling a V-Net, which minimized our proposed distance-penalized Cross-Entropy Loss (i.e. $\mathcal{L}= \frac{1}{N}\sum_{i=1}^{N}(1+\phi)\circ\sum_{j=1}^{K} CE$), where ϕ is the distance penalty term and CE the multiclass cross-entropy. We employed Adam Optimizer, learning rate 10-4 and due to computational constraints mini-batch included a single sample.

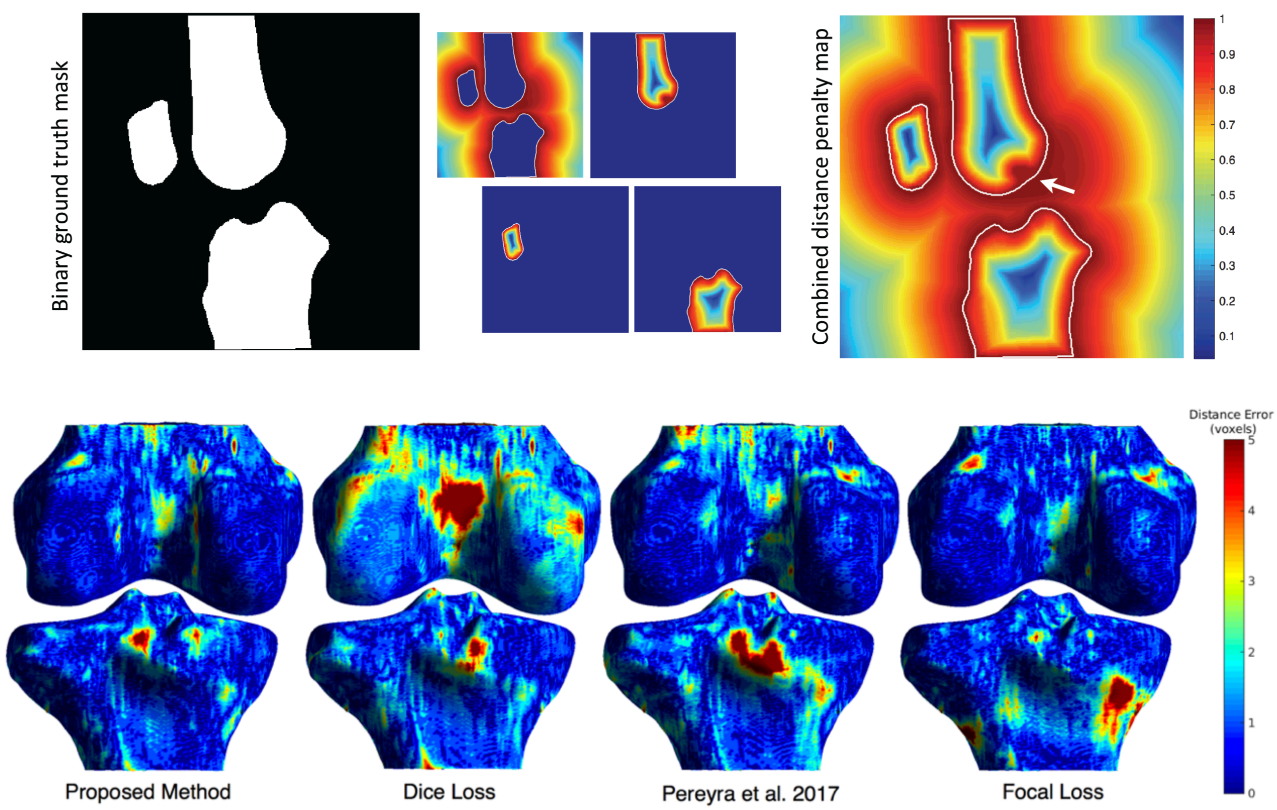

Figure 1: top) (L) Central sagittal slice showing binarized multi-class mask for distal femur, proximal tibia, and patella bones for a single training example. (C) Distance maps by compartment (normalized to [0,1]) before combining. (R) Combined multi-compartment 3D distance map derived from ground truth mask for loss penalty. Higher penalty is assigned to areas close to the bone boundaries (outlined in white). **bottom)** Posterior view of distal femur and proximal tibia for a representative test patient, predicted segmentation's absolute distance from ground truth. Proposed method produces high quality segmentations in areas particularly difficult to segment such as intercondylar notch or tibial condyle.

Results:
We investigated the effect of our penalty term, by comparing it to other successful losses, such as classical Dice loss, a penalization of confident outputs loss (PCOL) and focal loss (FL). Error maps in Figure 1-bottom show that our method produces high-quality segmentations even in regions with significant partial volume (intercondylar notch, tibial condyles). Quantitatively, the boundary dice score showed that our term significantly improved edge detection (0.29 ± 0.04) vs Dice loss (0.27 ± 0.05) vs PCOL (0.26 ± 0.03) vs focal-loss (0.27 ± 0.04). Global Dice Score showed that this superior performance is maintained globally (0.96 ± 0.08) vs Dice loss (0.96 ± 0.12) vs PCOL (0.95 ± 0.11) vs focal loss (0.95 ± 0.10).

Conclusion:
We argue that the segmentation ground truth provides additional information which, when appropriately utilized, can lead to improved segmentation of object boundaries difficult to precisely detect (especially in MRI where objects show diffused edges). In an extension of this work, we are currently coupling the semantic segmentation problem to that of voxel-wise distance maps regression in a multi-task learning setting. Arguably, this has the implicit effect of regularizing network training and improving the performance of each individual task. We ultimately aim to use our loss penalty term to improve the extraction of shape biomarkers and derive metrics to quantitatively evaluate the preservation of bone shape.

Highlights of Abstract:
We observed that guiding the network with a shape-aware loss function is a promising method to improve segmentation performance.