Knee Segmentation by Multiplanar Deep Learning Network – with data from OAI

Perslev, Mathias; Pai, Akshay Sadananda Uppinakudru; Igel, Christian; Dam, Erik Bjørnager

Publication date:
2018

Document version
Publisher's PDF, also known as Version of record

Knee Segmentation by Multiplanar Deep Learning Network – with data from OAI

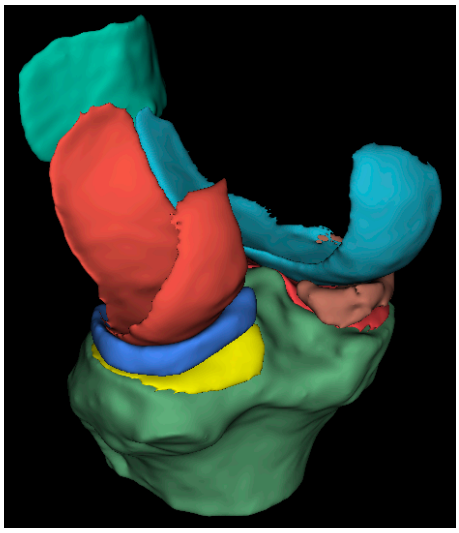
Mathias Perslev[1], Akshay Pai[1,2], Christian Igel[1], Erik B Dam[1,2]
[1] University of Copenhagen, Department of Computer Science
[2] Biomediq, Copenhagen, Denmark

PURPOSE: AI and particularly Deep Learning methodologies are receiving much attention. We validated a fully automatic convolutional neural network originally developed for brain MRI segmentation for the task of knee MRI segmentation.

METHODS: We analyzed the subset of 88 baseline knee MRIs from the Osteoarthritis Initiative (OAI) with semi-manual segmentations including patellar and medial/lateral tibial/femoral cartilages, medial/lateral menisci, and the Tibia. Using 44 scans; we trained a multi-planar convolutional neural network with a U-net inspired architecture. The remaining 44 scans were used for validation. We quantified segmentation performance by the Dice volume overlap for cartilages and menisci.

RESULTS:

Dice Volume Overlap (mean + std)		
Comp.	Training (n=44)	Validation (n=44)
Bone		
Tibia	0.98+0.00	0.98+0.00
Cartilages		
MT	0.85+0.05	0.83+0.08
LT	0.89+0.03	0.87+0.06
MF	0.88+0.02	0.84+0.05
LF	0.90+0.02	0.87+0.04
P	0.84+0.07	0.80+0.12
Menisci		
MM	0.84+0.05	0.80+0.08
LM	0.88+0.02	0.86+0.03



DISCUSSION: Deep Learning networks are potentially prone to overfitting and poor generalization. By using a network designed for brain MRI we eliminated **method overfitting**. Our evaluation showed similar performance on training and validation sets, albeit demonstrating slight **parameter overfitting**. Extending the training set with the corresponding 1 year scans and performing cross-validation (e.g. 5 fold CV) will both increase the training set size and very likely reduce this parameter overfitting. These results and previous work indicate that validation performance very close to the given training performance is feasible and that this is very likely very close to the limit of the information content in this cohort.

Segmentation time is approximately 2 minutes making the method applicable in clinical workflows. As demonstrated, our network is robust for the small, homogeneous data set. Further validation will evaluate the performance on larger, heterogeneous cohorts.

CONCLUSION: The study demonstrated the feasibility of accurate, fast, automatic knee MRI segmentation.

DISCLOSURE: Erik Dam is a shareholder of Biomediq.