

# Anatomical Priors for Weakly Supervised Osteophyte Detection and Localization in Hip X-rays

## Evaluating BoneFinder-Derived Guidance Under Image-Level Supervision

### 1. About this work

- Osteoarthritis is a common disease that can be assessed by observing osteophytes in X-ray images
- Models trained with weak labels could help with osteophyte detection
- Anatomical information is a form of weak supervision that may improve prediction and localization
- **Research question:** Can anatomical priors improve the spatial plausibility and robustness of weakly supervised osteophyte localization under limited image-level supervision?

### 2. Methodology

- The first layer of a ResNet18 architecture was modified to accept an additional channel with guidance information obtained from anatomical landmarks generated by the Bonefinder tool
- The analyzed models are: 1-channel (1ch), 2-channel (1ch+guidance), 3-channel (3ch) and 4-channel (3ch+guidance)

### 3. Results

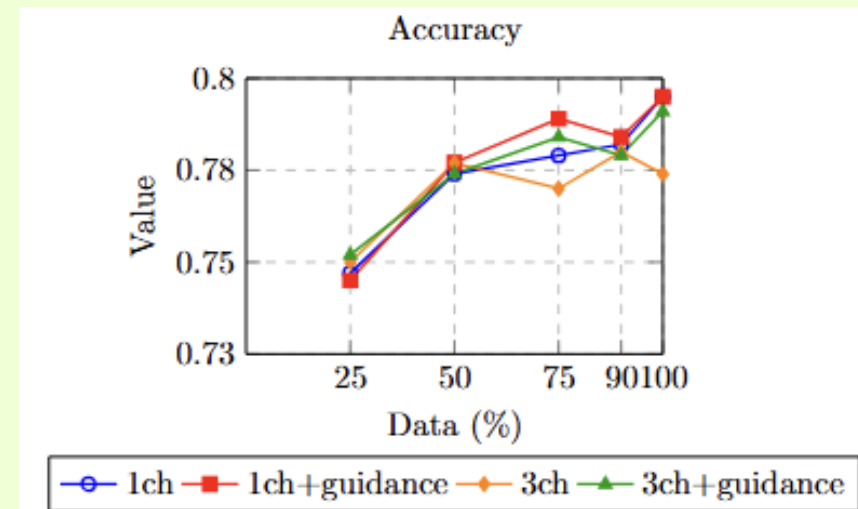


Figure 3: Accuracy across different training set sizes for the four models analyzed

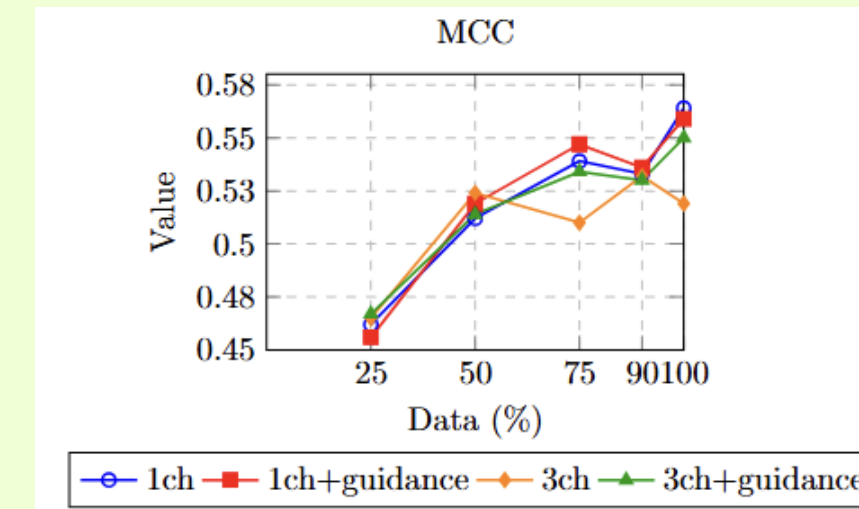


Figure 4: MCC across different training set sizes for the four models analyzed

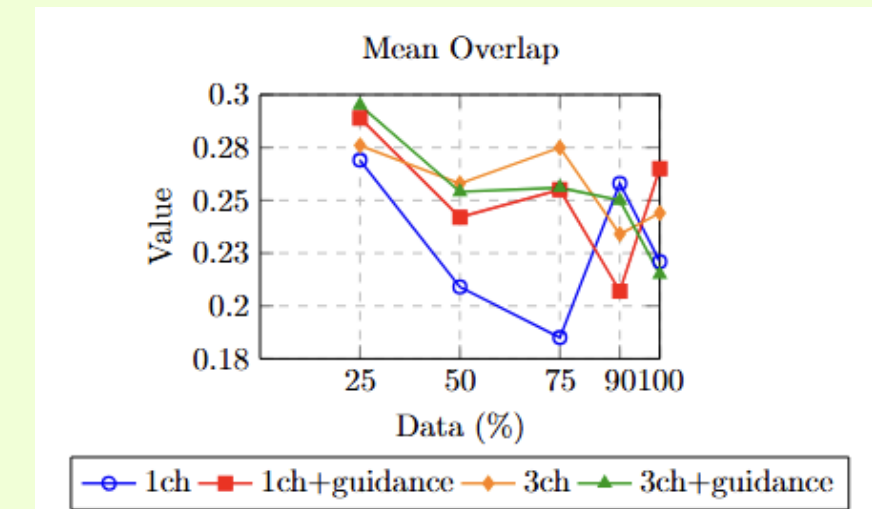
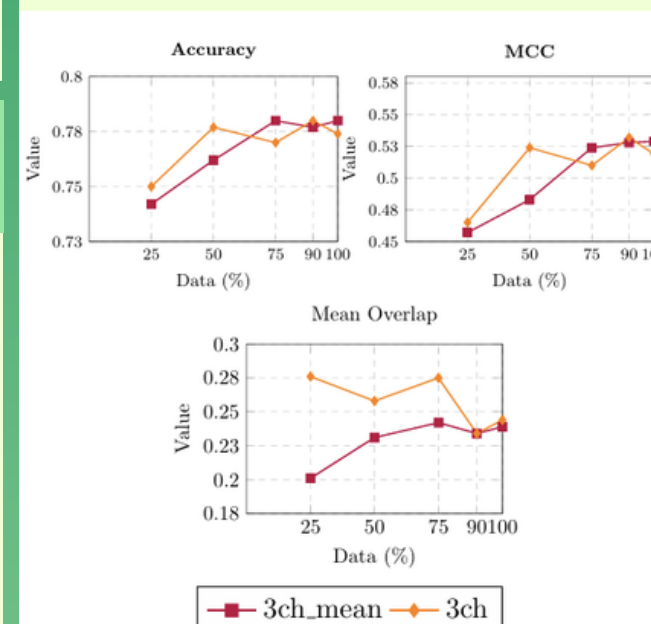


Figure 5: Overlap across different training set sizes for the four models analyzed

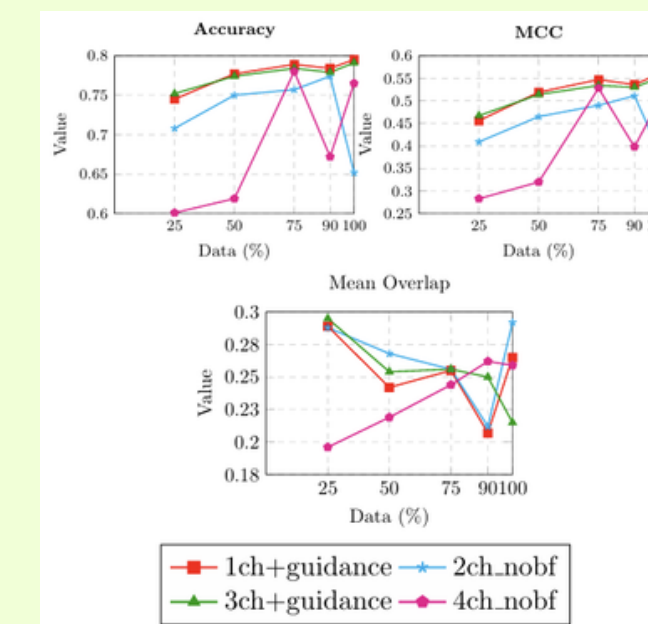
### Changing initialization



The initialization of the weights of the channels are the standard pretrained values (3ch) or the mean of these values for every channel (3ch\_mean).

Figure 6: Classification performance (Accuracy and MCC, top) and localization performance (Mean Overlap, bottom) across varying training set sizes for the 3ch and 3ch\_mean models.

### Removing guidance



Results obtained when we remove the guidance at test time (2ch\_nobf and 4ch\_nobf) for the analyzed models

Figure 7: Classification performance (Accuracy and MCC, top) and localization performance (Mean Overlap, bottom) across varying training set sizes for the 1ch+guidance, 2ch\_nobf, 3ch+guidance and 4ch\_nobf models.

### 4. Discussion and Conclusion

- Incorporating guidance generally improves classification performance with a less consistent impact on localization
- Since improvements are modest, this guidance cannot fully substitute for missing data
- Preserving pretrained RGB weights is more beneficial for localization than using an averaged representation
- Removing guidance suggests that guidance helps the model identify finer details, rather than changing which part of the body it looks at

### 5. Limitations and future work

- BoneFinder guidance does not represent the ground truth and in real-life scenarios it might not be available
- The region of interest was not medically chosen
- The model might learn to identify regions rather than osteophytes
- More experiments with different parameters should be done
- More exploration of ResNet architectures can be done

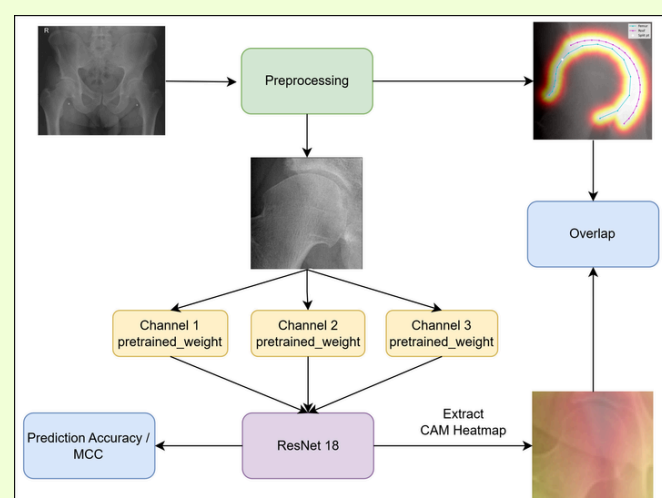


Figure 1: Example of the 3-channel ResNet pipeline.

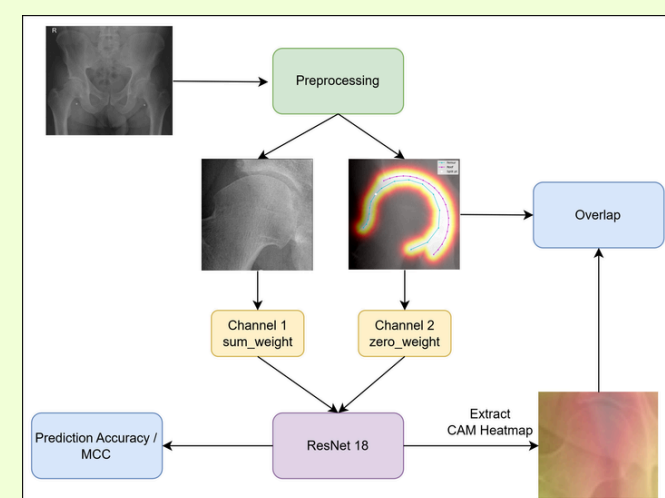


Figure 2: Example of the 2-channel ResNet pipeline.