Guiding Weakly Supervised Hip Osteophyte Detection Using Landmark-Based Anatomical Priors

CSE3000 Research Project · Author: Michel Chen · Responsible Professor: Jesse Krijthe · Supervisor: Gijs van Tulder

1) Weak Supervision Needs Spatial Guidance

- Osteophytes are bone protrusions in the hip.
- Weak supervision uses only image-level labels.
- Image-level labels do not tell the model where to look.
- Model may classify using irrelevant or correlated regions.
- Similar methods still use dense, pixel-level annotations.

Good classification does not guarantee anatomical localization.

2) Research Question:

How do sparse landmark-based anatomical priors affect **classification** and heatmap **localization** in weakly supervised hip osteophyte detection?

3) From landmarks to spatial penalty

Baseline model:

- CNN trained with image-level labels.
- Loss function: binary cross-entropy (L_{BCE}).

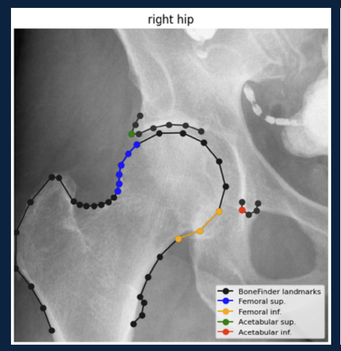
Prior-guided model:

- Selected landmarks define expected osteophyte zones.
- Landmark distances are converted into a penalty mask.
- Far-away CAM activation receives higher penalty:

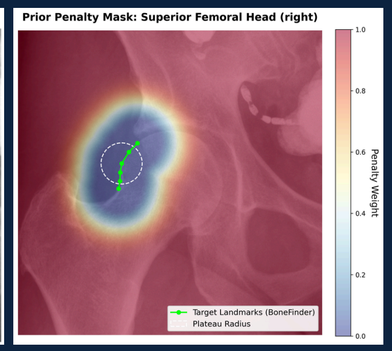
$$penalty = 1 - \exp\left(-\frac{\max(0, d-r)^2}{2\sigma^2}\right)$$

- d: distance to nearest target landmark
- r: no-penalty radius
- σ: smoothness of penalty
- Total loss: $\mathcal{L}_{\text{total}} = \mathcal{L}_{\text{BCE}} + \lambda\mathcal{L}_{\text{prior}}$
- λ: strength of spatial penalty

The model still uses image-level labels, but the penalty guides activation towards the anatomical target zone.



Coloured landmarks define the anatomical target zones.



Activation far from the target zone is penalized.

Setup: CHECK + OAI hip radiographs · ResNet-18 · 4 osteophyte targets · localization evaluated on positive labels

4) Result 1: AUC stays similar

Target	Baseline	Prior
Superior Femoral	0.92	0.92
Inferior Femoral	0.90	0.89
Superior Acetabular	0.79	0.80
Inferior Acetabular	0.87	0.86

AUC change: -0.01 to +0.01
Classification is preserved

The prior changes where the model looks, but does not add new diagnostic information to the image.

5) Result 2: Localization improves

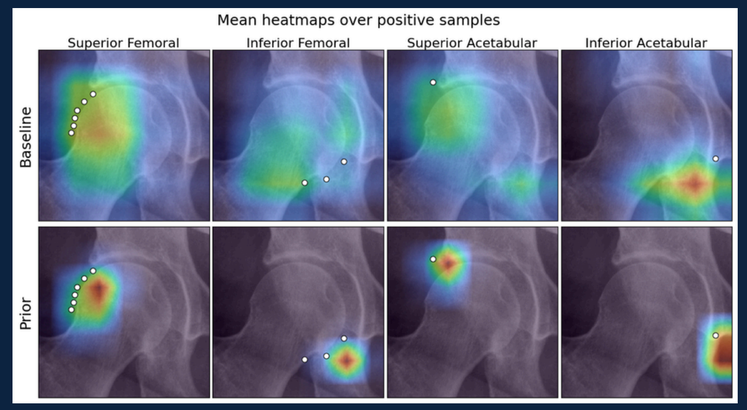
Target	Center-of-mass distance	Peak distance	Heatmap spread
Sup. Femoral	−23.9 px	−24.9 px	−24.5 px
Inf. Femoral	−28.3 px	−36.2 px	−33.3 px
Sup. Acetabular	−74.0 px	−73.1 px	−33.7 px
Inf. Acetabular	−44.8 px	−48.8 px	−30.0 px

Mean difference on positive target labels. Lower is better.

- All paired Wilcoxon tests are significant ($p < 0.001$).
- Largest improvement: superior acetabular localization, where baseline heatmaps were most displaced.

The prior consistently improves anatomical alignment.

6) Result 3: Heatmaps become more compact and closer to target zones



- Baseline: broad or displaced activation.
  - Co-occurring targets may allow proxy regions.
- Prior: activation closer to landmark-defined zones.
  - Prior penalty reduces target ambiguity.
  - More compact, but not exact localization.

The prior makes heatmaps more compact and regionally aligned, but does not provide exact localization.

7) Limitation: better alignment is not exact localization

- Prior depends on quality of the landmark points.
- Landmarks define approximate target zones, but true osteophytes may extend around them.
- Final feature map is only 7×7, so prior is spatially coarse.
- In some true positives, the prior lowers confidence.

The method improves regional alignment, not exact lesion segmentation.

8) Conclusion:

Landmark-based penalty masks substantially **improve** heatmap alignment and compactness while **preserving** classification performance