

## Supplementary

	number of patients	number of vertebrae	normal	grade			morphology		
				mild	moderate	severe	wedge	crush	concave
<i>train</i>	2481	9862	9319	227	222	94	309	229	5
<i>test</i>	438	1743	1643	33	47	20	47	2	51
<i>total</i>	2919	11605	10962	260	269	114	356	231	56

Table 1: *Demographic statistics.* Fractures and relevant properties of the annotated subset of the SUPERB cohort.

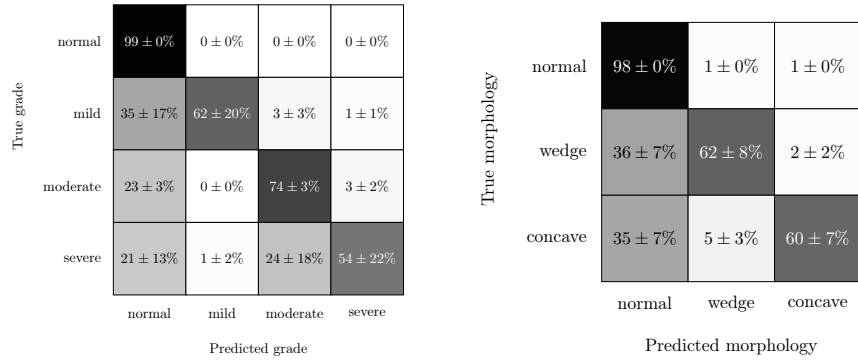


Fig. 1: *Confusion matrices.* Confusion matrices for the classification of each severity grade and morphology class, normalized by true classes.

	Detector model	Keypoints model
<i>learning rate</i>	$10^{-6}$	$5 \times 10^{-5}$
<i>weight decay</i>	$10^{-6}$	$5 \times 10^{-5}$
<i>batch size</i>	8	64
<i>equalization</i>	✓	✓
<i>inversion</i>	✓	✓
<i>rotation</i>	15%	20%
<i>flips</i>	✓	✓
<i>blur</i>	✓	✓
<i>random bbox scale</i>	✗	35%
<i>random bbox jitter</i>	✗	10 pixels
$\lambda_{iou}$	2	-
$\lambda_{\ell_1}$	5	-

Table 2: *Hyperparameters and training details.* Parameters used for training model components, chosen by grid search.

OVR target	$\mathcal{L}_{rle}$	$\mathcal{L}_{kps}$	$\mathcal{L}_{img}$	AUC	$F_1$	sensitivity	specificity
<i>normal</i>	✓			$0.96 \pm 0.01$	$0.48 \pm 0.02^*$	$0.90 \pm 0.03^*$	$0.89 \pm 0.01^*$
	✓	✓		$0.96 \pm 0.01$	$0.45 \pm 0.06^*$	$0.91 \pm 0.03^*$	$0.86 \pm 0.04^*$
	✓		✓	<b><math>0.98 \pm 0.01^*</math></b>	<b><math>0.71 \pm 0.06^*</math></b>	$0.91 \pm 0.04$	<b><math>0.96 \pm 0.01</math></b>
	✓	✓	✓	$0.92 \pm 0.04$	$0.69 \pm 0.08$	<b><math>0.98 \pm 0.01</math></b>	$0.95 \pm 0.02$
<i>wedge-like</i>	✓			$0.960 \pm 0.004^*$	$0.950 \pm 0.006$	$0.920 \pm 0.012$	$0.93 \pm 0.02$
	✓	✓		$0.95 \pm 0.01^*$	$0.960 \pm 0.001$	$0.920 \pm 0.002$	$0.90 \pm 0.03$
	✓		✓	$0.980 \pm 0.017$	<b><math>0.98 \pm 0.01</math></b>	<b><math>0.96 \pm 0.01</math></b>	$0.92 \pm 0.05$
	✓	✓	✓	<b><math>0.98 \pm 0.01</math></b>	$0.97 \pm 0.02$	$0.95 \pm 0.04$	<b><math>0.93 \pm 0.04</math></b>
<i>concave</i>	✓			$0.94 \pm 0.01^*$	$0.95 \pm 0.01^*$	$0.90 \pm 0.03^*$	$0.86 \pm 0.04^*$
	✓	✓		$0.93 \pm 0.01^*$	$0.95 \pm 0.01^*$	$0.92 \pm 0.03$	$0.86 \pm 0.03^*$
	✓		✓	$0.98 \pm 0.01$	$0.97 \pm 0.02$	$0.95 \pm 0.03$	$0.925 \pm 0.030$
	✓	✓	✓	<b><math>0.98 \pm 0.01</math></b>	<b><math>0.98 \pm 0.01</math></b>	<b><math>0.96 \pm 0.03</math></b>	<b><math>0.93 \pm 0.03</math></b>

Table 3: *Ablation study of loss components.* 5-fold comparison of the vertebra morphology classification on the test set ( $N_{vertebra} = 1743$ ).  $F_1$ , specificity and sensitivity computed at the Youden operating point. (\*) indicates significant difference to the full model using a Bonferroni  $t$ -test at  $\alpha = 0.05$ .