

1. The Statistics of Implants

There are 120 images with implants out of the 240 training images and 24 out of 60 testing images. We also calculated the evaluation metrics on the testing cases with and without implants separately (Table 1). It is shown that in both subgroups, the proposed modules improve the performance considerably. On those cases without metal implant, our method can correctly identify all vertebrae.

2. Clarification on Localization Error

The localization error is calculated only on the **correctly identified** vertebrae to eliminate the impact of the outliers. This is the convention on evaluating using this dataset by previous literature. In our ablation study, the id. rate consistently increases as the novel modules (the **leftmost** column in Table 1) are added. Nevertheless the localization errors are not always reduced when more vertebrae are correctly identified and then included in the calculation.

We also evaluated the localization errors for **all** vertebrae including outliers. The results are reported in Table 1 (Column “*incl. outlier*”). We observe the consistent reduction of the mean error as the novel modules are added. This result serves as a conclusive evidence of the effectiveness of the proposed modules.

3. Implausible Cases in the Prediction

We added the number of implausible cases in Table 1. The *base* and *base+rect* model produce 17 and 5 (out of 60) implausible results. With our anatomically-constrained optimization, all obtained results are physically plausible.

Table 1. Performance comparison in different settings

Model	Original (60)			Implant cases (24)			No implants (36)			incl. outlier		Implausible #
	Mean Error	Std	Id Rate	Mean Error	Std	Id Rate	Mean Error	Std	Id Rate	Mean Error	Std	
base	2.46	1.48	81.9	2.83	1.81	75.3	2.27	1.24	85.6	3.23	4.78	17
base+rect	2.54	1.68	81.4	2.89	2.06	74.9	2.37	1.42	85.1	3.32	4.61	5
base+rect+order	2.57	1.7	91.5	2.92	1.97	80.8	2.41	1.53	97.6	2.67	2.39	0
base+rect+optim	2.55	1.4	97.4	2.83	1.48	92.9	2.4	1.33	100.0	2.55	1.4	0
base+rect+order wo. λ	2.57	1.58	89.0	2.92	1.98	79.9	2.4	1.31	94.2	2.65	2.05	0
base+rect+optim wo. λ	2.55	1.39	94.4	2.83	1.48	92.9	2.39	1.3	95.2	2.59	1.63	0