For this task we used a classical transfer learning approach.

Starting from the given features we trained a backbone with two different regression heads.

One for the lumu regression and one for the gap.

The first one was the first to be trained. Since we have the most available data for it we can extract good features.

After the lumu regression training was completed, we froze the backbone and switched the head to the gap regression one.

We first only trained the head weights and after we got decent results we unfroze and refined the whole network.

When performing this transfer operations, batch normalization played an important role.