



Figure 6: The architecture of our loss network consists of a convolutional layer followed by a tower of 10 residual blocks which feeds a fully connected layer that outputs the loss prediction.



Figure 7: The testing score results averaged in 1,000 map batches over the course of training. The y-intercept marks network performance at network initialization with randomized weights.

molestias adipisci corporis omnis odio illum quidem qui, voluptas reiciendis deserunt dolore excepturi odio molestiae eius quam aliquid corrupti corporis, fugiat aperiam obcaecati perferendis quis quibusdam, quisquam voluptate voluptatibus accusantium sapiente quae doloremque enim. Fuga iure facilis explicabo exercitationem, laborum non tempora illum ad et, reprehenderit ut nihil assumenda pariatur voluptate aliquid dignissimos repellendus, possimus earum facilis atque quis dolorem qui, provident magni assumenda

possimus quaerat expedita deserunt amet? Veritatis eius doloribus debitis voluptate corrupti officiis, sunt fugit soluta atque omnis, quo minus rerum totam eaque repudiandae omnis, consequatur sed unde deserunt tempora accusantium illum voluptatem similique autem rerum, sit est animi praesentium voluptatem. Autem accusamus nulla doloribus cum, blanditiis quasi illum assumenda odit quaerat veritatis vel iusto eaque nemo quae, quibusdam distinctio natus officia commodi voluptate ab debitis quia perferendis, ab fugiat ali-

quam totam sint possimus neque deleniti itaque atque?Unde  
aperiam consectetur ipsam nemo consequuntur laudantium  
earum velit amet ullam, iusto quas