Appendix

A. Derivation of the Kullback-Leibler Divergence of Laplace Distributions

The Kullback-Leibler (KL) divergence between a probability distribution $q(x)$ and a reference distribution $p(x)$ is defined as follows:

$$\begin{aligned} & D(p(x)\|q(x)) \\ &= H(p(x), q(x)) - H(p(x)) \\ &= -\int_{-\infty}^{\infty} p(x)\log q(x)dx + \int_{-\infty}^{\infty} p(x)\log p(x)dx \end{aligned} \tag{33}$$

where $H(p(x))$ is the entropy of $p(x)$ and $H(p(x), q(x))$ is the cross entropy between $p(x)$ and $q(x)$. When both $p(x)$ and $q(x)$ are Laplace distributions,

$$p(x) = \frac{1}{2b_1}\exp\left(-\frac{|x-\mu_1|}{b_1}\right) \tag{34}$$

and

$$q(x) = \frac{1}{2b_2}\exp\left(-\frac{|x-\mu_2|}{b_2}\right), \tag{35}$$

the cross entropy becomes

$$\begin{aligned} & H(p(x), q(x)) \\ &= -\int_{-\infty}^{\infty} p(x)\log q(x)dx \\ &= \int_{-\infty}^{\infty} \frac{|x-\mu_2|}{2b_1b_2}\exp\left(-\frac{|x-\mu_1|}{b_1}\right)dx + \log(2b_2). \end{aligned} \tag{36}$$

To evaluate the integral, consider the case when $\mu_1 \geq \mu_2$,

$$\begin{aligned} & \int_{-\infty}^{\infty} \frac{|x-\mu_2|}{2b_1b_2}\exp\left(-\frac{|x-\mu_1|}{b_1}\right)dx \\ &= \int_{-\infty}^{\mu_2} \frac{\mu_2 - x}{2b_1b_2}\exp\left(-\frac{\mu_1 - x}{b_1}\right)dx \\ &+ \int_{\mu_2}^{\mu_1} \frac{x-\mu_2}{2b_1b_2}\exp\left(-\frac{\mu_1 - x}{b_1}\right)dx \\ &+ \int_{\mu_1}^{\infty} \frac{x-\mu_2}{2b_1b_2}\exp\left(-\frac{x-\mu_1}{b_1}\right)dx \end{aligned} \tag{37}$$

and when $\mu_1 < \mu_2$,

$$\begin{aligned} & \int_{-\infty}^{\infty} \frac{|x-\mu_2|}{2b_1b_2}\exp\left(-\frac{|x-\mu_1|}{b_1}\right)dx \\ &= \int_{-\infty}^{\mu_1} \frac{\mu_2 - x}{2b_1b_2}\exp\left(-\frac{\mu_1 - x}{b_1}\right)dx \\ &+ \int_{\mu_1}^{\mu_2} \frac{\mu_2 - x}{2b_1b_2}\exp\left(-\frac{x-\mu_1}{b_1}\right)dx \\ &+ \int_{\mu_2}^{\infty} \frac{x-\mu_2}{2b_1b_2}\exp\left(-\frac{x-\mu_1}{b_1}\right)dx. \end{aligned} \tag{38}$$

Evaluating each of the integrals produces the following result:

$$\begin{aligned} & \int_{-\infty}^{\infty} \frac{|x-\mu_2|}{2b_1b_2}\exp\left(-\frac{|x-\mu_1|}{b_1}\right)dx \\ &= \begin{cases} \frac{b_1\exp\left(-\frac{\mu_1-\mu_2}{b_1}\right)+(\mu_1-\mu_2)}{b_2}, & \mu_1 \geq \mu_2 \\ \frac{b_1\exp\left(-\frac{\mu_2-\mu_1}{b_1}\right)+(\mu_2-\mu_1)}{b_2}, & \mu_1 < \mu_2. \end{cases} \end{aligned} \tag{39}$$

Altogether, the cross entropy between two Laplace distribution is

$$\begin{aligned} & H(p(x), q(x)) \\ &= -\int_{-\infty}^{\infty} p(x)\log q(x)dx \\ &= \frac{b_1\exp\left(-\frac{|\mu_1-\mu_2|}{b_1}\right) + |\mu_1-\mu_2|}{b_2} + \log(2b_2) \end{aligned} \tag{40}$$

and the entropy of a Laplace distribution is

$$H(p(x)) = -\int_{-\infty}^{\infty} p(x)\log p(x)dx = 1 + \log(2b_1). \tag{41}$$

As a result, the KL divergence between two Laplace distributions is

$$\begin{aligned} & D(p(x)\|q(x)) \\ &= \frac{b_1\exp\left(-\frac{|\mu_1-\mu_2|}{b_1}\right) + |\mu_1-\mu_2|}{b_2} + \log\frac{b_2}{b_1} - 1. \end{aligned} \tag{42}$$

B. Proof of Differentiability

In Section 4.2, we utilize a second-order approximation of our proposed loss function (Equation (16)) about zero to illustrate its relationship with the Huber loss (Equation (4)). In this section, we will derive the first and second derivatives of our loss and prove their existence at zero.

Equation (16) can be written as follows:

$$D_{\alpha,\beta}(x) = \begin{cases} \frac{\alpha\exp\left(-\frac{x}{\alpha}\right)+x-\alpha}{\beta} & x \geq 0 \\ \frac{\alpha\exp\left(\frac{x}{\alpha}\right)-x-\alpha}{\beta} & x < 0. \end{cases} \tag{43}$$

Therefore, its first derivative is

$$\begin{aligned} D'_{\alpha,\beta}(x) &= \begin{cases} \frac{1-\exp\left(-\frac{x}{\alpha}\right)}{\beta} & x \geq 0 \\ -\frac{1-\exp\left(\frac{x}{\alpha}\right)}{\beta} & x < 0 \end{cases} \\ &= \frac{\operatorname{sgn}(x)}{\beta}\left(1-\exp\left(-\frac{|x|}{\alpha}\right)\right), \end{aligned} \tag{44}$$

and its second derivative is

$$D''_{\alpha,\beta}(x) = \begin{cases} \frac{\exp\left(-\frac{x}{\alpha}\right)}{\alpha\beta} & x \geq 0 \\ \frac{\exp\left(\frac{x}{\alpha}\right)}{\alpha\beta} & x < 0 \end{cases} = \frac{1}{\alpha\beta}\exp\left(-\frac{|x|}{\alpha}\right). \tag{45}$$

To prove the existence of the derivatives at $x = 0$, we need to show that both $D_{\alpha,\beta}(x)$ and $D'_{\alpha,\beta}(x)$ are differentiable at $x = 0$. The derivative of any function $f(x)$ at $x = a$ is defined as follows:

$$f'(a) = \lim_{x \to a} \frac{f(x) - f(a)}{x - a}. \tag{46}$$

The function $f(x)$ is said to be differentiable at $x = a$ when the limit exists. To prove our claims, we will rely heavily on the following well-known identity:

$$\lim_{x \to 0} (1 + x)^{\frac{1}{x}} = e. \tag{47}$$

Claim 1. *The following function is differentiable at $x = 0$:*

$$D_{\alpha,\beta}(x) = \frac{\alpha \exp\left(-\frac{|x|}{\alpha}\right) + |x| - \alpha}{\beta} \tag{48}$$

when $\alpha > 0$ and $\beta > 0$.

Proof. The derivative of $D_{\alpha,\beta}(x)$ at $x = 0$ is defined as

$$\begin{aligned} D'_{\alpha,\beta}(0) &= \lim_{x \to 0} \frac{\alpha \exp\left(-\frac{|x|}{\alpha}\right) + |x| - \alpha}{\beta x} \\ &= \lim_{x \to 0} \frac{\alpha\left(\exp\left(-\frac{|x|}{\alpha}\right) - 1\right)}{\beta x} + \lim_{x \to 0} \frac{|x|}{\beta x}. \end{aligned} \tag{49}$$

The right limit of the first term is equal to the following:

$$\lim_{x \to 0^+} \frac{\alpha\left(\exp\left(-\frac{x}{\alpha}\right) - 1\right)}{\beta x} = \lim_{u \to 0^+} -\frac{u}{\beta \log(u+1)} \tag{50}$$

where we substitute $u = \exp(-x/\alpha) - 1$ in for x; therefore, $x = -\alpha \log(u + 1)$ and $u \to 0$ as $x \to 0$. Utilizing Equation (47), the limit becomes

$$\begin{aligned} \lim_{u \to 0^+} -\frac{u}{\beta \log(u+1)} &= \lim_{u \to 0^+} -\frac{1}{\beta \log(u+1)^{\frac{1}{u}}} \\ &= -\frac{1}{\beta \log\left(\lim_{u \to 0^+}(u+1)^{\frac{1}{u}}\right)} \\ &= -\frac{1}{\beta}. \end{aligned} \tag{51}$$

Similarly, for the left limit of the first term,

$$\lim_{x \to 0^-} \frac{\alpha\left(\exp\left(\frac{x}{\alpha}\right) - 1\right)}{\beta x} = \lim_{v \to 0^-} \frac{v}{\beta \log(v+1)} = \frac{1}{\beta} \tag{52}$$

where we substitute $v = \exp(x/\alpha) - 1$ in for x; as a result, $x = \alpha \log(v + 1)$ and $v \to 0$ as $x \to 0$. Furthermore, the right limit of the second term is

$$\lim_{x \to 0^+} \frac{|x|}{\beta x} = \lim_{x \to 0^+} \frac{x}{\beta x} = \frac{1}{\beta}, \tag{53}$$

and the left limit is

$$\lim_{x \to 0^-} \frac{|x|}{\beta x} = \lim_{x \to 0^-} -\frac{x}{\beta x} = -\frac{1}{\beta}. \tag{54}$$

By adding both terms together, both sides of the limit become zero, which means the limit exists and proves $D_{\alpha,\beta}(x)$ is differentiable at $x = 0$. □

Claim 2. *The following function is differentiable at $x = 0$:*

$$D'_{\alpha,\beta}(x) = \frac{\operatorname{sgn}(x)}{\beta}\left(1 - \exp\left(-\frac{|x|}{\alpha}\right)\right) \tag{55}$$

when $\alpha > 0$ and $\beta > 0$.

Proof. The derivative of $D'_{\alpha,\beta}(x)$ at $x = 0$ is defined as

$$D''_{\alpha,\beta}(0) = \lim_{x \to 0} \frac{\operatorname{sgn}(x)\left(1 - \exp\left(-\frac{|x|}{\alpha}\right)\right)}{\beta x}. \tag{56}$$

The right limit is equal to the following:

$$\lim_{x \to 0^+} \frac{1 - \exp\left(-\frac{x}{\alpha}\right)}{\beta x} = \lim_{u \to 0^+} \frac{u}{\alpha\beta \log(u+1)} \tag{57}$$

where we substitute $u = \exp(-x/\alpha) - 1$ in for x; therefore, $x = -\alpha \log(u + 1)$ and $u \to 0$ as $x \to 0$. Utilizing Equation (47), the limit becomes

$$\begin{aligned} \lim_{u \to 0^+} \frac{u}{\alpha\beta \log(u+1)} &= \lim_{u \to 0^+} \frac{1}{\alpha\beta \log(u+1)^{\frac{1}{u}}} \\ &= \frac{1}{\alpha\beta \log\left(\lim_{u \to 0^+}(u+1)^{\frac{1}{u}}\right)} \\ &= \frac{1}{\alpha\beta} \end{aligned} \tag{58}$$

Similarly, for the left limit,

$$\lim_{x \to 0^-} -\frac{1 - \exp\left(\frac{x}{\alpha}\right)}{\beta x} = \lim_{v \to 0^-} \frac{v}{\alpha\beta \log(v+1)} = \frac{1}{\alpha\beta} \tag{59}$$

where we substitute $v = \exp(x/\alpha) - 1$ in for x; as a result, $x = \alpha \log(v + 1)$ and $v \to 0$ as $x \to 0$. Therefore, both sides of the limit are equal, which means the limit exists and proves $D'_{\alpha,\beta}(x)$ is differentiable at $x = 0$. □

C. Proof of Inequalities

In Section 4.2, we state that the Huber loss, $H_\alpha(x)$, is bounded below by $D_{\alpha,1/\alpha}(x)$ and above by $D_{\alpha/2,1/\alpha}(x)$, and the bounds are tight. In this section, we prove these claims. Since the loss functions are symmetric about $x = 0$, it is sufficient to prove only when $x \geq 0$.

Claim 3. *The following inequality holds for all $x \in \mathbb{R}$:*

$$H_\alpha(x) - D_{\alpha,1/\alpha}(x) \geq 0 \tag{60}$$

Proof. When $0 \leq x \leq \alpha$, the inequality is

$$\frac{1}{2}x^2 - \alpha x + \alpha^2 - \alpha^2 \exp\left(-\frac{x}{\alpha}\right) \geq 0 \tag{61}$$

and it becomes

$$\frac{1}{2}\alpha^2 - \alpha^2 \exp\left(-\frac{x}{\alpha}\right) \geq 0 \tag{62}$$

when $x \geq \alpha$. The inequalities can be simplified by substituting $y = x/\alpha$ and dividing by α^2. As a result, we now need to prove

$$f_1(y) = \frac{1}{2}y^2 - y + 1 - \exp(-y) \geq 0 \tag{63}$$

when $0 \leq y \leq 1$, and

$$f_2(y) = \frac{1}{2} - \exp(-y) \geq 0 \tag{64}$$

when $y \geq 1$. Equation (63) is a well-known inequality and can be proven by utilizing the mean value theorem. The first and second derivative of $f_1(y)$ are

$$f_1'(y) = y - 1 + \exp(-y) \tag{65}$$

and

$$f_1''(y) = 1 - \exp(-y). \tag{66}$$

When $y \geq 0$, $f_1'(y) \geq 0$ since $f_1''(y) \geq 0$ and $f_1'(0) = 0$; likewise, $f_1(y) \geq 0$ for the same reason, $f_1'(y) \geq 0$ and $f_1(0) = 0$. The proof of the second inequality follows directly from the first. From Equation (63), we know that $\exp(-1) \leq 1/2$; therefore, $f_2(y) \geq 0$ when $y \geq 1$ since $\exp(-y)$ is monotonically decreasing. □

Claim 4. *The following inequality holds for all $x \in \mathbb{R}$:*

$$D_{\alpha/2,1/\alpha}(x) - H_\alpha(x) \geq 0 \tag{67}$$

Proof. The inequality is equal to

$$\frac{\alpha^2}{2}\exp\left(-\frac{2}{\alpha}x\right) + \alpha x - \frac{\alpha^2}{2} - \frac{1}{2}x^2 \geq 0 \tag{68}$$

when $0 \leq x \leq \alpha$, and it is

$$\frac{\alpha^2}{2}\exp\left(-\frac{2}{\alpha}x\right) \geq 0 \tag{69}$$

when $x \geq \alpha$. Again, the inequalities can be simplified by substituting $y = 2x/\alpha$ and dividing by $\alpha^2/2$, which results in the following inequalities:

$$f_3(y) = \exp(-y) + y - 1 - \frac{1}{4}y^2 \geq 0 \tag{70}$$

when $0 \leq y \leq 2$, and

$$f_4(y) = \exp(-y) \geq 0 \tag{71}$$

when $y \geq 2$. The second inequality, $f_4(y) \geq 0$, clearly holds for all $y \in \mathbb{R}$; whereas, the first inequality, $f_3(y) \geq 0$, is less obvious. The first and second derivative of $f_3(y)$ are

$$f_3'(y) = -\exp(-y) + 1 - \frac{1}{2}y \tag{72}$$

and

$$f_3''(y) = \exp(-y) - \frac{1}{2}. \tag{73}$$

At $y = 0$, $f_3'(0) = 0$ and $f_3''(0) = 1/2 > 0$, and at $y = 2$, $f_3'(2) = -\exp(-2) < 0$. Since $f_3''(y)$ has a single root, $f_3'(y)$ can have at most two roots by Rolle's theorem. Therefore, there exists a unique value, $0 < y_0 < 2$, where $f_3'(y_0) = 0$, and on the interval $0 \leq y \leq y_0$, $f_3'(y) \geq 0$. Moreover, by the mean value theorem, $f_3(y) \geq 0$ on that interval, $0 \leq y \leq y_0$, since $f_3'(y) \geq 0$ and $f_3(0) = 0$. Note that $f_3'(y) \leq 0$, or equivalently

$$\exp(-y) \geq 1 - \frac{1}{2}y \tag{74}$$

on the interval $y_0 \leq y \leq 2$. Consequently, to complete the proof of $f_3(y) \geq 0$, we just need to show that

$$1 - \frac{1}{2}y \geq 1 - y + \frac{1}{4}y^2 \tag{75}$$

or correspondingly

$$f_5(y) = -\frac{1}{4}y^2 + \frac{1}{2}y \geq 0 \tag{76}$$

when $y_0 \leq y \leq 2$. The roots of $f_5(y)$ are at $y = 0$ and $y = 2$, since $f_5(1) = 1/4 > 0$, $f_5(y) \geq 0$ on the interval $0 \leq y \leq 2$. □

Claim 5. *For all $x \in \mathbb{R}$, $H_\alpha(x)$ is tightly bounded between $D_{\alpha,1/\alpha}(x)$ and $D_{\alpha/2,1/\alpha}(x)$. Therefore, the inequalities*

$$D_{\alpha,1/\alpha}(x) \leq D_{\alpha_1,\beta_1}(x) \leq H_\alpha(x) \tag{77}$$

and

$$H_\alpha(x) \leq D_{\alpha_2,\beta_2}(x) \leq D_{\alpha/2,1/\alpha}(x) \tag{78}$$

hold only, for all $x \in \mathbb{R}$, when $\alpha_1 = \alpha$, $\alpha_2 = \alpha/2$, and $\beta_1 = \beta_2 = 1/\alpha$.

Proof. The inequalities are equivalent to

$$D_{\alpha,1/\alpha}(x) - H_\alpha(x) \leq D_{\alpha_1,\beta_1}(x) - H_\alpha(x) \leq 0 \tag{79}$$

and

$$0 \leq D_{\alpha_2,\beta_2}(x) - H_\alpha(x) \leq D_{\alpha/2,1/\alpha}(x) - H_\alpha(x). \tag{80}$$

As x goes to infinity,

$$\lim_{x\to\infty} D_{\alpha/2,1/\alpha}(x) - H_\alpha(x) = 0 \quad (81)$$

$$\lim_{x\to\infty} D_{\alpha,1/\alpha}(x) - H_\alpha(x) = -\frac{1}{2}\alpha^2 \quad (82)$$

and

$$\lim_{x\to\infty} D_{\alpha_*,\beta_*}(x) - H_\alpha(x) = \begin{cases} \infty, & \beta_* < \frac{1}{\alpha} \\ \alpha\left(\frac{1}{2}\alpha - \alpha_*\right), & \beta_* = \frac{1}{\alpha} \\ -\infty, & \beta_* > \frac{1}{\alpha}. \end{cases} \quad (83)$$

For the inequalities to hold in the limit, β_* must equal $1/\alpha$ regardless of the value of α_*, α_2 must equal $\alpha/2$, and α_1 must be between $\alpha/2$ and α, inclusively. Now, we need to demonstrate that there exists an $x \in \mathbb{R}$ where

$$D_{\alpha_1,1/\alpha}(x) - H_\alpha(x) > 0 \quad (84)$$

when $\alpha/2 < \alpha_1 < \alpha$. The inequality is equal to

$$\alpha\left(\alpha_1 \exp\left(-\frac{x}{\alpha_1}\right) + x - \alpha_1\right) - \frac{1}{2}x^2 > 0 \quad (85)$$

when $0 \le x \le \alpha$. To simplify the inequality, let us set $\alpha_1 = \alpha/\gamma$, substitute $y = \gamma x/\alpha$, and divide by α^2/γ where $1 < \gamma < 2$, which results in the following inequality:

$$f_6(y) = \exp(-y) + y - 1 - \frac{1}{2\gamma}y^2 > 0 \quad (86)$$

when $0 \le y \le \gamma$. The first and second derivative of $f_6(y)$ are

$$f_6'(y) = -\exp(-y) + 1 - \frac{1}{\gamma}y \quad (87)$$

and

$$f_6''(y) = \exp(-y) - \frac{1}{\gamma}. \quad (88)$$

At $y = 0$, $f_6'(0) = 0$ and $f_6''(0) = 1 - 1/\gamma > 0$ for $1 < \gamma < 2$, and at $y = \gamma$, $f_6'(\gamma) = -\exp(-\gamma) < 0$. Like before, by Rolle's theorem, $f_6'(y)$ can have at most two roots since $f_6''(y)$ has a single root. Therefore, there exists a unique value, $0 < y_0 < \gamma$, where $f_6'(y_0) = 0$, and on the interval $0 < y < y_0$, $f_6'(y) > 0$. Again, by the mean value theorem, $f_6(y) > 0$ on that interval, $0 < y < y_0$, since $f_6'(y) > 0$ and $f_6(0) = 0$. Therefore, α_1 must equal α for the original inequalities to hold. □

D. Experimental Validation of Target Approximation

In Section 6, we claim the target width and target height can be approximated with the percentage change between the anchor and the ground-truth. To validate the approximation, we train the Faster R-CNN model with the following targets:

$$t_w^* = \frac{w^*}{w_a} - 1 \quad (89)$$

and

$$t_h^* = \frac{h^*}{h_a} - 1. \quad (90)$$

No other changes were made to the implementation. Refer to Section 7.1, for details on the training and evaluation procedure. The results of the experiment are shown in Table 5. Only a very slight degradation in performance is observed by replacing the targets with its approximation, which we believe validates our use of the approximation in our interpretation of the loss functions.

Table 5: Target Performance

Target	Mean Average Precision (mAP) @		
	0.5 IoU	0.75 IoU	0.5-0.95 IoU
Original	**44.7**	**23.1**	**23.8**
Approximation	44.6	23.0	23.7

E. Evaluation of Proposed Loss Function

Although our goal is to understand the Huber loss and not to replace it, for the sake of completion, we demonstrate that replacing the Huber loss with our proposed loss function produces comparable results. As mentioned in Section 4.2, minimizing the loss function $D_{\alpha/2,1/\alpha}$ is equivalent to minimizing an upper-bound on the Huber loss H_α. Therefore, for this experiment we simply replace H_α with $D_{\alpha/2,1/\alpha}$ with no other modifications to Faster R-CNN. Refer to Section 7.1, for details on the training and evaluation procedure. The results are shown in Table 6. The performance of the loss functions are nearly identical, which is expected due to the similarity of the functions.

Table 6: Loss Function Performance

Loss Function	Mean Average Precision (mAP) @		
	0.5 IoU	0.75 IoU	0.5-0.95 IoU
H_α	**44.7**	23.1	**23.8**
$D_{\alpha/2,1/\alpha}$	**44.7**	**23.3**	**23.8**