1 Algorithm

1.1 General Description

The target of this model is to find out a more reliable score based on Bayesian optimization for each bounding boxes given by faster-rcnn based on the output scores, as well as the relative position, number and size as prior information from training data set.

1.2 Notation

Faster-rcnn is a local detector based on Convolutional neural network. For each testing image, it provides 100 bounding boxes for each category, with its position and predicted score.
1.We define the position of each bounding box as $B = (x, y, h, w)$, where (x, y) is the center coordination of each bounding box, h is the height, w is the width of the bounding box.
2.The the position of ground truth bounding box is also defined in the same wayas $Gt = (x, y, h, w)$.
3.The intersection of union, which measures the overlap between two bounding boxes, or bounding boxes and ground truth bounding box, is defined as $IoU = \frac{Area \quad of \quad Overlap}{Area \quad of \quad Union}$.
4.$\boldsymbol{t}$ is the indicator sequence with length 100, which is corresponding to the 100 bounding boxes for category table given by faster-rcnn. It is used to indicating the prediction if for each bounding box, if it really contains a table or not. $\boldsymbol{c}$ is the indicator sequence with length 100 for category chair.
5.$\boldsymbol{s_t}$ and $\boldsymbol{s_c}$ is the predicted score sequence of table and chair from faster-rcnn.
6. $\boldsymbol{\eta_t}$ and $\boldsymbol{\eta_c}$ is the bounding box sequence of table and chair, which means for each entry, it contains the coordination information $B = (x, y, h, w)$ of this bounding box.
7.γ are the parameters for score distributions; $\boldsymbol{\theta}$ are the parameters for relative coordination distribution; $\boldsymbol{\tau}$ are the parameters for size distribution and $\boldsymbol{\lambda}$ are the parameters for number distribution.

1.3 Model

The target of model is to maximize the conditional probability:

$$\boldsymbol{c}, \boldsymbol{t} = \arg\max_{\boldsymbol{c}, \boldsymbol{t}} p(\boldsymbol{c}, \boldsymbol{t} | \boldsymbol{s_t}, \boldsymbol{s_c}, \boldsymbol{\eta_t}, \boldsymbol{\eta_c}, \gamma, \boldsymbol{\theta}, \boldsymbol{\tau}, \boldsymbol{\lambda})$$

However, it is very difficult to maximize this joint distribution straightforwardly. Therefore we use MCMC method to update instead.

1.Initialization: Each indicator in $\boldsymbol{c}, \boldsymbol{t}$ is 1 except the following conditions:

- IoU between that position and previous position are greater than 0.7.

- Score is smaller than 0.1.

- The previous total number of 1's in the chair sequence already reaches 8, or in the table sequence already reaches 5.

2.Update each chair score with other tables and chairs fixed:

Sort the chair sequence from current highest score to the lowest score. For each chair in the sequence, calculate the following conditional probability:
$p(c_i = 1|\boldsymbol{c_{-i}}, \boldsymbol{t}, \boldsymbol{s_t}, \boldsymbol{s_c}, \boldsymbol{\eta_t}, \boldsymbol{\eta_c}, \gamma, \boldsymbol{\theta}, \boldsymbol{\tau}, \boldsymbol{\lambda})$

$$= \frac{p(\boldsymbol{s_t}, \boldsymbol{s_c}|c_i = 1, \boldsymbol{c_{-i}}, \boldsymbol{t}, \boldsymbol{\eta_t}, \boldsymbol{\eta_c}, \gamma, \boldsymbol{\theta}, \boldsymbol{\tau}, \boldsymbol{\lambda})p(c_i = 1|\boldsymbol{c_{-i}}, \boldsymbol{t}, \boldsymbol{\eta_t}, \boldsymbol{\eta_c}, \gamma, \boldsymbol{\theta}, \boldsymbol{\tau}, \boldsymbol{\lambda})}{\sum\limits_{c_i=0,1} p(\boldsymbol{s_t}, \boldsymbol{s_c}|c_i, \boldsymbol{c_{-i}}, \boldsymbol{t}, \boldsymbol{\eta_t}, \boldsymbol{\eta_c}, \gamma, \boldsymbol{\theta}, \boldsymbol{\tau}, \boldsymbol{\lambda})p(c_i|\boldsymbol{c_{-i}}, \boldsymbol{t}, \boldsymbol{\eta_t}, \boldsymbol{\eta_c}, \gamma, \boldsymbol{\theta}, \boldsymbol{\tau}, \boldsymbol{\lambda})} \tag{1}$$

Now we split the equation into several parts based on different prior information.

$$p(\boldsymbol{s_t}, \boldsymbol{s_c}|c_i = 1, \boldsymbol{c_{-i}}, \boldsymbol{t}, \boldsymbol{\eta_t}, \boldsymbol{\eta_c}, \gamma, \boldsymbol{\theta}, \boldsymbol{\tau}, \boldsymbol{\lambda}) = p(\boldsymbol{s_t}, \boldsymbol{s_c}|c_i = 1, \boldsymbol{c_{-i}}, \boldsymbol{t}, \gamma) = p(\boldsymbol{s_t}|\boldsymbol{t}, \gamma)p(\boldsymbol{s_{c_i}}|\boldsymbol{c_i}, \gamma)p(\boldsymbol{s_{c_{-i}}}|\boldsymbol{c_{-i}}, \gamma) \tag{2}$$

The above part only uses the score prior information.
$p(c_i = 1|\boldsymbol{c_{-i}}, \boldsymbol{t}, \boldsymbol{\eta_t}, \boldsymbol{\eta_c}, \gamma, \boldsymbol{\theta}, \boldsymbol{\tau}, \boldsymbol{\lambda})$

$$= \frac{p(\boldsymbol{t}|c_i = 1, \boldsymbol{c_{-i}}, \boldsymbol{\eta_t}, \boldsymbol{\eta_c}, \boldsymbol{\theta}, \boldsymbol{\tau}, \boldsymbol{\lambda})p(c_i = 1|\boldsymbol{c_{-i}}, \boldsymbol{\eta_t}, \boldsymbol{\eta_c}, \boldsymbol{\theta}, \boldsymbol{\tau}, \boldsymbol{\lambda})}{p(\boldsymbol{t}|\boldsymbol{c_{-i}}, \boldsymbol{\eta_t}, \boldsymbol{\eta_c}, \boldsymbol{\theta}, \boldsymbol{\tau}, \boldsymbol{\lambda})} \tag{3}$$

For the above equation, we reduce the probability into a form conditioning on fixed numbers, which become distributions depending on number, relative position and size priors. The denominator is only about $\boldsymbol{t}$, so it will be canceled in the equation (1) by nominator and denominator in the end.

$$\begin{aligned} &p(\boldsymbol{t}|c_i = 1, \boldsymbol{c_{-i}}, \boldsymbol{\eta_t}, \boldsymbol{\eta_c}, \boldsymbol{\theta}, \boldsymbol{\tau}, \boldsymbol{\lambda}) \\ &\quad = p(\boldsymbol{t}|\|\boldsymbol{t}\|_1, c_i = 1, \boldsymbol{c_{-i}}, \boldsymbol{\eta_t}, \boldsymbol{\eta_c}, \boldsymbol{\theta}, \boldsymbol{\tau}, \boldsymbol{\lambda})p(\|\boldsymbol{t}\|_1|c_i = 1, \boldsymbol{c_{-i}}, \boldsymbol{\eta_t}, \boldsymbol{\eta_c}, \boldsymbol{\theta}, \boldsymbol{\tau}, \boldsymbol{\lambda}) \end{aligned} \tag{4}$$

$$\begin{aligned} p(\|\boldsymbol{t}\|_1|c_i = 1, \boldsymbol{c_{-i}}, \boldsymbol{\eta_t}, \boldsymbol{\eta_c}, \boldsymbol{\theta}, \boldsymbol{\tau}, \boldsymbol{\lambda}) & \\ &= p(\|\boldsymbol{t}\|_1|c_i = 1, \boldsymbol{c_{-i}}, \boldsymbol{\lambda}) \\ &= p(\|\boldsymbol{t}\|_1|\|\boldsymbol{c_{-i}}\|_1 + 1, \boldsymbol{\lambda}) \\ &= \frac{p(\|\boldsymbol{c_{-i}}\|_1 + 1|\|\boldsymbol{t}\|_1, \boldsymbol{\lambda})p(\|\boldsymbol{t}\|_1|\boldsymbol{\lambda})}{\sum\limits_{\|\boldsymbol{t}\|_1} p(\|\boldsymbol{c_{-i}}\|_1 + 1|\|\boldsymbol{t}\|_1, \boldsymbol{\lambda})p(\|\boldsymbol{t}\|_1|\boldsymbol{\lambda})} \end{aligned} \tag{5}$$

Equation (5) only uses the number information, with constrain on number of tables.

suppose σ is permutation of $(t_{r_1}, t_{r_2}, t_{r_3} \cdots t_{r_k})$,with $t_{r_1} \cdots t_{r_k} = 1$, and $t_{(1\cdots n)\setminus(r_1\cdots r_k)} = 0$.

$$p(\boldsymbol{t}|\|\boldsymbol{t}\|_1, c_i = 1, \boldsymbol{c_{-i}}, \boldsymbol{\eta_t}, \boldsymbol{\eta_c}, \boldsymbol{\theta}, \boldsymbol{\tau}, \boldsymbol{\lambda}) = \sum_{j=1\cdots k!} p_{\sigma(j)}(\boldsymbol{t}|\|\boldsymbol{t}\|_1, c_i = 1, \boldsymbol{c_{-i}}, \boldsymbol{\eta_t}, \boldsymbol{\eta_c}, \boldsymbol{\theta}, \boldsymbol{\tau}, \boldsymbol{\lambda}) \tag{6}$$

Each term in the sum is the probability that the 1's in the sequence $\sigma(i)$ are selected in order when the number of 1's is $\|\boldsymbol{t}\|_1$.
To explain this calculation process specifically, let the sequence of permutation $\sigma(j)$ be $(t_{r_{j_1}}, \cdots, t_{r_{j_k}})$.

$$p_{\sigma(j)}(\boldsymbol{c}|\|\boldsymbol{c}\|_1, t_i = 1, \boldsymbol{t_{-i}}, \boldsymbol{\eta_t}, \boldsymbol{\eta_c}, \boldsymbol{\theta}, \boldsymbol{\tau}, \boldsymbol{\lambda}) = \prod_{m \in (1,\cdots,k)} \frac{\dfrac{p(t_{r_{j_m}}|\boldsymbol{c})}{\sum_{h \in (1,\cdots,n)} p(t_h|\boldsymbol{c})}}{1 - \sum_{q<m} \dfrac{p(t_{r_{j_q}}|\boldsymbol{c})}{\sum_{h \in (1,\cdots,n)} p(t_h|\boldsymbol{c})}}$$

Where $p(t_h|\boldsymbol{c}) = max_{u \in (r_1,\cdots,r_l)} p(t_h|c_u)$, with $c_{r_1} \cdots c_{r_l} = 1$, and $c_{(1\cdots n)\backslash(r_1 \cdots r_l)} = 0$.

$$p(c_i = 1|\boldsymbol{c_{-i}}, \boldsymbol{\eta_t}, \boldsymbol{\eta_c}, \boldsymbol{\theta}, \boldsymbol{\tau}, \boldsymbol{\lambda})$$
$$= \frac{p(c_i = 1, \boldsymbol{c_{-i}}|\|\boldsymbol{c_{-i}}\|_1 + 1, \boldsymbol{\eta_t}, \boldsymbol{\eta_c}, \boldsymbol{\theta}, \boldsymbol{\tau}, \boldsymbol{\lambda}) p(\|\boldsymbol{c_{-i}}\|_1 + 1|\boldsymbol{\eta_t}, \boldsymbol{\eta_c}, \boldsymbol{\theta}, \boldsymbol{\tau}, \boldsymbol{\lambda})}{p(\boldsymbol{c_{-i}}|\boldsymbol{\eta_t}, \boldsymbol{\eta_c}, \boldsymbol{\theta}, \boldsymbol{\tau}, \boldsymbol{\lambda})} \quad (7)$$

suppose σ is the permutation of $(c_{r_1}, c_{r_2}, c_{r_3} \cdots t_{r_l})$, in which $c_{r_1} \cdots c_{r_l} = 1$ under the restriction $c_i = 1$, and $c_{(1\cdots n)\backslash(r_1 \cdots r_l)} = 0$.

$$p(c_i = 1, \boldsymbol{c_{-i}}|\|\boldsymbol{c_{-i}}\|_1 + 1, \boldsymbol{\eta_t}, \boldsymbol{\eta_c}, \boldsymbol{\theta}, \boldsymbol{\tau}, \boldsymbol{\lambda}) = \sum_{j=1\cdots l!} p_{\sigma(j)}(c_i = 1, \boldsymbol{c_{-i}}|\|\boldsymbol{c_{-i}}\|_1 + 1, \boldsymbol{\eta_t}, \boldsymbol{\eta_c}, \boldsymbol{\theta}, \boldsymbol{\tau}, \boldsymbol{\lambda}) \quad (8)$$

let the sequence of permutation $\sigma(j)$ be $(c_{r_{j_1}}, \cdots, c_{r_{j_l}})$.

$$p_{\sigma(j)}(c_i = 1, \boldsymbol{c_{-i}}|\|\boldsymbol{c_{-i}}\|_1 + 1, \boldsymbol{\eta_t}, \boldsymbol{\eta_c}, \boldsymbol{\theta}, \boldsymbol{\tau}, \boldsymbol{\lambda}) = \prod_{m \in (1,\cdots,l)} \frac{\dfrac{p(c_{r_{j_m}})}{\sum_{h \in (1,\cdots,n)} p(c_h)}}{1 - \sum_{q<m} \dfrac{p(c_{r_{j_q}})}{\sum_{h \in (1,\cdots,n)} p(c_h)}}$$

$$p(\|\boldsymbol{c_{-i}}\|_1 + 1|\boldsymbol{\eta_t}, \boldsymbol{\eta_c}, \boldsymbol{\theta}, \boldsymbol{\tau}, \boldsymbol{\lambda}) = \sum_{\|\boldsymbol{t}\|} p(\|\boldsymbol{c_{-i}}\|_1 + 1|\|\boldsymbol{t}\|_1, \boldsymbol{\lambda}) p(\|\boldsymbol{t}\|_1|\boldsymbol{\lambda}) \quad (9)$$

Now we can rewrite and calculate the updating equation (1) using the above equations.

3.The updating for table scores based on other tables and chairs fixed is almost in the same process:

Sort the table sequence from current highest score to the lowest score. For each table in the sequence, calculate the following conditional probability:
$p(t_i = 1|\boldsymbol{t_{-i}}, \boldsymbol{c}, \boldsymbol{s_t}, \boldsymbol{s_c}, \boldsymbol{\eta_t}, \boldsymbol{\eta_c}, \boldsymbol{\gamma}, \boldsymbol{\theta}, \boldsymbol{\tau}, \boldsymbol{\lambda})$

$$= \frac{p(\boldsymbol{s_t}, \boldsymbol{s_c}|t_i = 1, \boldsymbol{t_{-i}}, \boldsymbol{c}, \boldsymbol{\eta_t}, \boldsymbol{\eta_c}, \boldsymbol{\gamma}, \boldsymbol{\theta}, \boldsymbol{\tau}, \boldsymbol{\lambda})p(t_i = 1|\boldsymbol{t_{-i}}, \boldsymbol{c}, \boldsymbol{\eta_t}, \boldsymbol{\eta_c}, \boldsymbol{\gamma}, \boldsymbol{\theta}, \boldsymbol{\tau}, \boldsymbol{\lambda})}{\sum_{t_i=0,1} p(\boldsymbol{s_t}, \boldsymbol{s_c}|t_i, \boldsymbol{t_{-i}}, \boldsymbol{c}, \boldsymbol{\eta_t}, \boldsymbol{\eta_c}, \boldsymbol{\gamma}, \boldsymbol{\theta}, \boldsymbol{\tau}, \boldsymbol{\lambda})p(t_i|\boldsymbol{t_{-i}}, \boldsymbol{c}, \boldsymbol{\eta_t}, \boldsymbol{\eta_c}, \boldsymbol{\gamma}, \boldsymbol{\theta}, \boldsymbol{\tau}, \boldsymbol{\lambda})} \tag{10}$$

$$p(\boldsymbol{s_t}, \boldsymbol{s_c}|t_i = 1, \boldsymbol{t_{-i}}, \boldsymbol{c}, \boldsymbol{\eta_t}, \boldsymbol{\eta_c}, \boldsymbol{\gamma}, \boldsymbol{\theta}, \boldsymbol{\tau}, \boldsymbol{\lambda}) = p(\boldsymbol{s_t}, \boldsymbol{s_c}|t_i = 1, \boldsymbol{t_{-i}}, \boldsymbol{c}, \gamma) = p(\boldsymbol{s_c}|\boldsymbol{c}, \gamma)p(\boldsymbol{s_{t_i}}|\boldsymbol{t_i}, \gamma)p(\boldsymbol{s_{t_{-i}}}|\boldsymbol{t_{-i}}, \gamma) \tag{11}$$

$$\begin{aligned} &p(t_i = 1|\boldsymbol{t_{-i}}, \boldsymbol{c}, \boldsymbol{\eta_t}, \boldsymbol{\eta_c}, \boldsymbol{\gamma}, \boldsymbol{\theta}, \boldsymbol{\tau}, \boldsymbol{\lambda}) \\ &\quad = \frac{p(\boldsymbol{c}|t_i = 1, \boldsymbol{t_{-i}}, \boldsymbol{\eta_t}, \boldsymbol{\eta_c}, \boldsymbol{\theta}, \boldsymbol{\tau}, \boldsymbol{\lambda})p(t_i = 1|\boldsymbol{t_{-i}}, \boldsymbol{\eta_t}, \boldsymbol{\eta_c}, \boldsymbol{\theta}, \boldsymbol{\tau}, \boldsymbol{\lambda})}{p(\boldsymbol{c}|\boldsymbol{t_{-i}}, \boldsymbol{\eta_t}, \boldsymbol{\eta_c}, \boldsymbol{\theta}, \boldsymbol{\tau}, \boldsymbol{\lambda})} \end{aligned} \tag{12}$$

Similarly, the denominator will be canceled at the end.

$$\begin{aligned} &p(\boldsymbol{c}|t_i = 1, \boldsymbol{t_{-i}}, \boldsymbol{\eta_t}, \boldsymbol{\eta_c}, \boldsymbol{\theta}, \boldsymbol{\tau}, \boldsymbol{\lambda}) \\ &= p(\boldsymbol{c}|\|\boldsymbol{c}\|_1, t_i = 1, \boldsymbol{t_{-i}}, \boldsymbol{\eta_t}, \boldsymbol{\eta_c}, \boldsymbol{\theta}, \boldsymbol{\tau}, \boldsymbol{\lambda})p(\|\boldsymbol{c}\|_1|t_i = 1, \boldsymbol{t_{-i}}, \boldsymbol{\eta_t}, \boldsymbol{\eta_c}, \boldsymbol{\theta}, \boldsymbol{\tau}, \boldsymbol{\lambda}) \end{aligned} \tag{13}$$

$$p(\|\boldsymbol{c}\|_1|t_i = 1, \boldsymbol{t_{-i}}, \boldsymbol{\eta_t}, \boldsymbol{\eta_c}, \boldsymbol{\theta}, \boldsymbol{\tau}, \boldsymbol{\lambda}) = p(\|\boldsymbol{c}\|_1|t_i = 1, \boldsymbol{t_{-i}}, \boldsymbol{\lambda}) = p(\|\boldsymbol{c}\|_1|\|\boldsymbol{t_{-i}}\|_1 + 1, \boldsymbol{\lambda}) \tag{14}$$

Equation (5) only uses the number information, with constrain on number of chairs.

suppose σ is permutation of $(c_{r_1}, c_{r_2}, c_{r_3} \cdots c_{r_k})$,with $c_{r_1} \cdots c_{r_k} = 1$, and $c_{(1 \cdots n) \setminus (r_1 \cdots r_k)} = 0$.

$$p(\boldsymbol{c}|\|\boldsymbol{c}\|_1, t_i = 1, \boldsymbol{t_{-i}}, \boldsymbol{\eta_t}, \boldsymbol{\eta_c}, \boldsymbol{\theta}, \boldsymbol{\tau}, \boldsymbol{\lambda}) = \sum_{j=1 \cdots k!} p_{\sigma(j)}(\boldsymbol{c}|\|\boldsymbol{c}\|_1, t_i = 1, \boldsymbol{t_{-i}}, \boldsymbol{\eta_t}, \boldsymbol{\eta_c}, \boldsymbol{\theta}, \boldsymbol{\tau}, \boldsymbol{\lambda}) \tag{15}$$

$$\begin{aligned} &p(t_i = 1|\boldsymbol{t_{-i}}, \boldsymbol{\eta_t}, \boldsymbol{\eta_c}, \boldsymbol{\theta}, \boldsymbol{\tau}, \boldsymbol{\lambda}) \\ &= \frac{p(t_i = 1, \boldsymbol{t_{-i}}|\|\boldsymbol{t_{-i}}\|_1 + 1, \boldsymbol{\eta_t}, \boldsymbol{\eta_c}, \boldsymbol{\theta}, \boldsymbol{\tau}, \boldsymbol{\lambda})p(\|\boldsymbol{t_{-i}}\|_1 + 1|\boldsymbol{\eta_t}, \boldsymbol{\eta_c}, \boldsymbol{\theta}, \boldsymbol{\tau}, \boldsymbol{\lambda})}{p(\boldsymbol{t_{-i}}|\boldsymbol{\eta_t}, \boldsymbol{\eta_c}, \boldsymbol{\theta}, \boldsymbol{\tau}, \boldsymbol{\lambda})} \end{aligned} \tag{16}$$

Now the equation (16) is only about table sequences. Suppose σ is the permutation of $(t_{r_1}, t_{r_2}, t_{r_3} \cdots t_{r_l})$, in which $t_{r_1} \cdots t_{r_l} = 1$ under the restriction $t_i = 1$, and $t_{(1 \cdots n) \setminus (r_1 \cdots r_l)} = 0$.

$$p(t_i = 1, \boldsymbol{t_{-i}}|\|\boldsymbol{t_{-i}}\|_1 + 1, \boldsymbol{\eta_t}, \boldsymbol{\eta_c}, \boldsymbol{\theta}, \boldsymbol{\tau}, \boldsymbol{\lambda}) = \sum_{j=1 \cdots l!} p_{\sigma(j)}(t_i = 1, \boldsymbol{t_{-i}}|\|\boldsymbol{t_{-i}}\|_1 + 1, \boldsymbol{\eta_t}, \boldsymbol{\eta_c}, \boldsymbol{\theta}, \boldsymbol{\tau}, \boldsymbol{\lambda}) \tag{17}$$

$p(\|\boldsymbol{t_{-i}}\|_1 + 1|\boldsymbol{\eta_t}, \boldsymbol{\eta_c}, \boldsymbol{\theta}, \boldsymbol{\tau}, \boldsymbol{\lambda})$ can be easily calculated.
In the same way, we can rewrite and calculate the updating equation using the above equations.

1.4 Extreme Case

However, if $\boldsymbol{t}$ or $\boldsymbol{c}$ is 0's sequence when updating, it is impossible to calculate the updating equation, since there is no information about relative position. In particular, $p(\boldsymbol{t}|\|\boldsymbol{t}\|_1, \boldsymbol{c}, \boldsymbol{\eta_t}, \boldsymbol{\eta_c}, \boldsymbol{\theta}, \boldsymbol{\tau}, \boldsymbol{\lambda})$ and $p(\boldsymbol{c}|\|\boldsymbol{c}\|_1, \boldsymbol{t}, \boldsymbol{\eta_t}, \boldsymbol{\eta_c}, \boldsymbol{\theta}, \boldsymbol{\tau}, \boldsymbol{\lambda})$ cannot be calculated. Therefore, we assume when $\boldsymbol{t}$ or $\boldsymbol{c}$ is 0's sequence, this part of probability will keep the same in the nominator and denominator, thus can be canceled during calculation.

2 Implementation Details

2.1 Estimating distributions and parameters

There are 4 hyper-parameter vectors that need to be estimated: $\boldsymbol{\gamma}, \boldsymbol{\theta}, \boldsymbol{\tau}, \boldsymbol{\lambda}$.

$\gamma = (\gamma_{\boldsymbol{t_0}}, \gamma_{\boldsymbol{t_1}}, \gamma_{\boldsymbol{c_0}}, \gamma_{\boldsymbol{c_1}})$,where $\boldsymbol{\gamma_{t_1}}$ are parameters for faster-rcnn score distributions of those bounding boxes indicating table and really containing table, while $\boldsymbol{\gamma_{t_0}}$ are parameters for score distributions of those bounding boxes indicating table and but actually without table. Same explanation for $(\gamma_{\boldsymbol{c_0}}, \gamma_{\boldsymbol{c_1}})$. Here we assume the faster-rcnn output score are gamma distributions. We estimate the parameters based on output scores of training data from faster-rcnn.

$\boldsymbol{\theta}$ are relative position parameters. Given a bounding box containing chair and another bounding box containing table, we measure the relative position by 3 statistics: $(a, b, c) = (\frac{x_c - x_t}{r_t}, \frac{y_c - y_t}{r_t}, \frac{r_c}{r_t} - 1)$. Where (x_c, y_c, x_t, y_t) are coordinate of center points for bounding boxes that indicating containing chair and table. We assume (a, b, c) follows a multivariate Gaussian distribution, $\boldsymbol{\theta} = (\mu_a, \sigma_a, \mu_b, \sigma_b, \mu_c, \sigma_c)$. Then $a, b, c \sim N(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ Where $\boldsymbol{\Sigma}$ is the diagonal covariance matrix. We also estimate them based on training data.

$\boldsymbol{\tau}$ are vectors of parameters for size distributions, there are also 4 distributions:size of bounding boxes that containing chair,containing table, without chair but indicating there is a chair, without table but indicating there is a table.
The distributions are estimated using kernel density estimation.

$\boldsymbol{\lambda} = (\lambda_t, \boldsymbol{\lambda_{c|t}})$, which measures the general number of tables in a random image, and number of chairs given a certain number of tables in an image. We estimate those parameters by fitting a Poisson distribution.