**Equations:**

**Image Features:**

**GRU Features:**

**Feature Fusion Similarity:**

**Stacked Cross Attention weights:**

**Weighted Sum of Safety Rule vectors:**

**Cosine similarity between Image and Safety:**

**Final Matching Score (Pooling Function)**

**Triplet Loss:**

**EIOU Loss:**

**Soft-Non-Maximum Suppression:**