# Supplementary Material: A Step Forward Towards Trustworthy Risk-Aware Facial Retrieval (RA-FR)*

Muhammad Emmad Siddiqui[1[0009−0005−1575−8409]] and Muhammad Rafi[1[0000−0002−3673−5979]]

National University of Computer and Emerging Sciences, Karachi, Pakistan
https://www.nu.edu.pk/

## 1 Proof of Finite-Sample Risk Guarantee

In this section, we provide the theoretical justification for the risk control mechanism described in the main paper. We demonstrate that the RA-FR Controller guarantees that the population risk $\rho(\hat{\kappa})$ remains below the user-specified tolerance $\alpha$ with high probability $1 - \delta$.

### 1.1 Problem Setup

Let $\mathcal{D}_{cal} = \{(X_i^Q, Y_i)\}_{i=1}^n$ be a held-out calibration dataset of size $n$, drawn i.i.d. from the joint distribution $\mathcal{P}$. For a given calibration parameter $\kappa \in \mathbb{R}^+$, the dynamic set size for a query $X_i^Q$ is given by the adapter function:

$$K_i(\kappa) = \lceil \kappa \cdot \Phi[f_u(X_i^Q)] \rceil, \tag{1}$$

where $\Phi[f_u(\cdot)]$ is the normalized uncertainty score in $[0, 1]$.

We define the binary loss for the $i$-th calibration point as:

$$\ell_i(\kappa) = \mathbb{I}(Y_i \notin \mathcal{R}_\kappa(X_i^Q)), \tag{2}$$

where $\mathcal{R}_\kappa(X_i^Q)$ is the set of top-$K_i(\kappa)$ retrieved images. The loss is 1 if the true match is missed, and 0 otherwise.

The **Empirical Risk** on the calibration set is defined as:

$$\hat{\rho}(\kappa) = \frac{1}{n} \sum_{i=1}^{n} \ell_i(\kappa). \tag{3}$$

### 1.2 Risk Control Theorem

**Theorem 1.** *Let $\hat{\kappa}$ be the calibration parameter computed by:*

$$\hat{\kappa} = \inf \left\{ \kappa \in \mathbb{R}^+ : \hat{\rho}(\kappa) + \sqrt{\frac{\ln(1/\delta)}{2n}} \leq \alpha \right\}. \tag{4}$$

* Code available at: https://github.com/MuhammadEmmadSiddiqui/RA-FR

*Then, the probability that the true population risk exceeds $\alpha$ is bounded by $\delta$:*

$$\Pr(\rho(\hat{\kappa}) > \alpha) \leq \delta. \tag{5}$$

*Proof.* The losses $\ell_i(\kappa)$ are i.i.d. random variables bounded in $[0, 1]$. We utilize **Hoeffding's Inequality** to bound the deviation of the empirical risk from the true risk. For any fixed $\kappa$, Hoeffding's inequality states:

$$\Pr\left(\rho(\kappa) - \hat{\rho}(\kappa) \geq \epsilon\right) \leq \exp(-2n\epsilon^2). \tag{6}$$

Setting the right-hand side to $\delta$ and solving for $\epsilon$:

$$\epsilon = \sqrt{\frac{\ln(1/\delta)}{2n}}. \tag{7}$$

Thus, with probability at least $1 - \delta$, the true risk is bounded by the empirical risk plus the complexity term:

$$\rho(\kappa) \leq \hat{\rho}(\kappa) + \sqrt{\frac{\ln(1/\delta)}{2n}}. \tag{8}$$

The algorithm selects $\hat{\kappa}$ specifically to satisfy the condition that this upper bound is less than or equal to $\alpha$. Therefore, if the probabilistic bound holds (which occurs with probability $1 - \delta$), then $\rho(\hat{\kappa}) \leq \alpha$.

### 1.3 Monotonicity Efficient Search

The validity of efficiently finding $\hat{\kappa}$ relies on the monotonicity of the risk function.

**Lemma 1.** *The empirical risk $\hat{\rho}(\kappa)$ is a monotonically non-increasing function of $\kappa$.*

*Proof.* Let $\kappa_1 < \kappa_2$. From the adapter equation, $K_i(\kappa_1) \leq K_i(\kappa_2)$. Since the retrieval function $R(X^Q)$ returns ranked candidates, a larger $K$ implies a superset of candidates: $\mathcal{R}_{\kappa_1}(X^Q) \subseteq \mathcal{R}_{\kappa_2}(X^Q)$. If the true match $Y$ is present in the smaller set, it must be present in the larger set. Thus, the loss cannot increase: $\ell_i(\kappa_1) \geq \ell_i(\kappa_2)$. Averaging over $n$ samples, $\hat{\rho}(\kappa_1) \geq \hat{\rho}(\kappa_2)$.

This monotonicity allows us to find $\hat{\kappa}$ efficiently using a linear scan or binary search over a discretized grid of $\kappa$ values.

## 2 Algorithms

We present the pseudo-code for the Calibration phase (RA-FR Controller) and the Inference phase (RA-FR Adapter).

**Algorithm 1** RA-FR Controller (Calibration)

**Require:** Calibration Set $\mathcal{D}_{cal} = \{(X_i, Y_i)\}_{i=1}^{n}$
**Require:** User tolerance $\alpha \in (0, 1)$, Failure prob $\delta \in (0, 1)$
**Require:** Pre-trained Feature Extractor $fe$, Uncertainty Head $f_u$
**Compute Hoeffding Slack:**
$\epsilon \leftarrow \sqrt{\frac{\ln(1/\delta)}{2n}}$
**Initialize Search Grid:**
$\Lambda \leftarrow [0, \Delta, 2\Delta, \ldots, \kappa_{max}]$ ▷ Discretized grid for $\kappa$
**Evaluate Risk:**
**for** $\kappa$ in $\Lambda$ **do**
  $L_{sum} \leftarrow 0$
  **for** $i = 1$ to $n$ **do**
    $u_i \leftarrow \Phi[f_u(X_i)]$ ▷ Get normalized uncertainty
    $K_i \leftarrow \lceil \kappa \cdot u_i \rceil$ ▷ Calculate set size
    Retrieve top-$K_i$ set $\mathcal{R}_i$ using $fe(X_i)$
    **if** $Y_i \notin \mathcal{R}_i$ **then**
      $L_{sum} \leftarrow L_{sum} + 1$
    **end if**
  **end for**
  $\hat{\rho} \leftarrow L_{sum}/n$ ▷ Empirical Risk
  $UCB \leftarrow \hat{\rho} + \epsilon$ ▷ Upper Confidence Bound
  **if** $UCB \leq \alpha$ **then**
    **return** $\hat{\kappa} \leftarrow \kappa$ ▷ Found minimal safe $\kappa$
  **end if**
**end for**
**return** $\kappa_{max}$ ▷ Default if target not reached

**Algorithm 2** RA-FR Adapter (Inference)

**Require:** Test Query $X_{test}$
**Require:** Calibrated parameter $\hat{\kappa}$
**Require:** Database $D$
$feat \leftarrow fe(X_{test})$
$unc \leftarrow f_u(X_{test})$
$u_{norm} \leftarrow \Phi[unc]$
$K_{opt} \leftarrow \lceil \hat{\kappa} \cdot u_{norm} \rceil$
Calculate distances $d(feat, z)$ for all $z \in D$
Rank candidates by distance
**return** Top-$K_{opt}$ candidates