

Dissertation (2020)

CHAPTER 8. BIAS IN FACE RECOGNITION

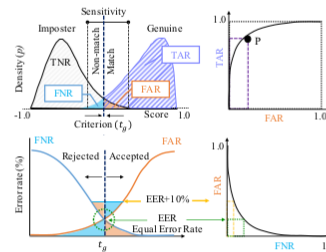


Figure 8.1: **Depiction of the biometrics.** The SDM shows the sensitivity related to a single threshold t_g (top-left). The area to the right of the threshold considers all accepted pairs, both correctly and incorrectly predicted. True Acceptance Rate (TAR) as a function of False Acceptance Rate (FAR) is a common way to report ratings for given false-rates (top-right). Equally common in FR is the trade-off between false-negative rate (FNR) and FAR (bottom-left and bottom-right).

that leads to skewed results across demographics and other attributes— are determined using a held-out set. In other words, the threshold for the FR *Matching* module is set according to the desired target: in research the threshold that yields that highest accuracy on the validation set; in practice the threshold is determined by the value that yields the desired rate, which depends on the use-case. For example, in the use-case where FR is used to enable entry via access control may have a smaller threshold that will realize fewer *genuine* samples than FR used for tagging photos per software recommendation—the falsely accepted instances of the latter use-case will have minimal, if any, negative effects. An important note to consider for the concept of having a held-out set is that it typically shares the same distribution with the test set, meaning it favors the same demographics as the held-out set had. That skew (*i.e.*, the difference in the performance of an algorithm of particular demographics) is our definition of bias. A key question is: *is FR too biased, or not?*

Now, provided two or more faces features encoded by a CNN, a distance (or similarity score) s must be learned such to act as a decision boundary to separate the genuine pairs score from the imposters score. Ideally, genuine and imposter scores would be completely separable. However, this is not the case in practice. It is this score-threshold (*i.e.*, θ) that determines whether or not the pair should be accepted. The implications are for faces features: to be assumed as the same, genuine