

Robust re-identification using randomness and statistical learning: Quo vadis

Michele Nappi, Harry Wechsler

The re-identification problem is to match objects across multiple but possibly disjoint fields of view for the purpose of sequential authentication over space and time. Detection and seeding for initialization do not presume known identity and allow for re-identification of objects and/or faces whose identity might remain unknown. Specific functionalities involved in re-identification include clustering and selection, recognition-by-parts, anomaly and change detection, sampling and tracking, fast indexing and search, sensitivity analysis, and their integration for the purpose of identity management. As re-identification processes data streams and involves change detection and on-line adaptation three complementary statistical

learning frameworks, driven by randomness for the purpose of robust prediction, are advanced here to support the functionalities listed earlier and their combination thereof. The intertwined learning frameworks employed are those of (a) semi-supervised learning (SSL); (b) transduction; and (c) conformal prediction. The overall architecture proposed is data-driven and modular, on one side, and discriminative and progressive, on the other side. The architecture is built around autonomic computing and W5+. Autonomic computing or self-management provides for closed-loop control. W5+ answers questions related to What data to consider for sampling and collection, When to capture the data and from Where, and How to best process the data. The Who (is) query is about identity for biometrics, and the Why question for explanation purposes. The challenge addressed throughout is that of evidence-based management to progressively collect and add value to data in order to generate knowledge that leads to purposeful and gainful action including active learning for the overall purpose of re-identification. A venue for future research includes adversarial learning when re-identification is possibly “distracted” using deliberate corrupt information.