УДК 004.8
Стець О.А.
Тернопільський національний технічний університет імені Івана Пулюя

АНАЛІЗ ДОМЕН-УЗАГАЛЬНЕНИХ МЕТОДІВ ВИЯВЛЕННЯ ПІДМІНИ ОБЛИЧ

Stets O.

DOMAIN-GENERALIZED FACE SPOOFING DETECTION METHODS ANALYSIS

Виявлення підміни обличь – це завдання запобігання атаки, ціль якої полягає у фальшивій верифікації обличчя за допомогою фото, відео, маски чи іншої заміни обличчя авторизованої особи. Приклади таких атак включають:

- «Друкована атака» – зловмисник використовує друковане або цифрове зображення іншої людини, яким підміняється зображення реальної автентифікованої персони.
- «Атака з відео» – складніший спосіб шахрайства, який зазвичай потребує відео обличчя жертви. Цей підхід гарантує, що поведінка та рухи обличчя виглядають більш природно порівняно з першим типом.
- «Атака 3D-маски» – під час атаки цього типу як інструмент підробки використовують маску. На додаток до природних рухів обличчя, це дозволяє обдурити й інші рівні захисту, такі як датчики глибини.

Тип атаки, а також особливості оточення, обладнання та сфери використання формують домен атаки. Методи захисту від підміни обличчя досягли відмінної ефективності в конкретних доменах, але залежність від домену є серйозною проблемою на шляху до масового впровадження таких методів у реальні сценарії використання. Тому розробка домен-узагальненого методу виявлення підміни обличчя є актуальним завданням, яке дозволить вирішити цілий ряд проблем безпеки при автентифікації.

Для обробки доменно-узагальнених невідомих атак в дослідженнях останніх місяців було представлено декілька новітніх методів. Метод DGUA-FAS складається з екстрактора ознак на основі перетворювача і синтетичного генератора зразків невідомих атак. Генератор імітує невідомі зразки атак, що допомагає в навчанні екстрактора функцій. Платформа ATR-FAS окрім домен-узагальненості та захисту від невідомих типів атак також надає захист від підміни обличчя під світловим спалахом. Вона містить подвійний шлюзовий модуль, що складається з шлюзу розпізнавання типу атаки типу та шлюзу формування кадрової уваги. Розробники методу FLIP вперше показали, що ініціалізація трансформаторів зору попередньо навченими мультимодальними ваговими коефіцієнтами покращує можливість узагальнення, що відповідає можливостям нульової передачі попередньо навчених моделей візуальної мови.

Такі розробки суттєво покращують гнучкість моделей та наближають впровадження цих методів у промислове використання, що демонструє їхню перспективність та потребу у таких дослідженнях у майбутньому.

Література:

1. George A., Marcel S. Idiap Research Institute. Deep Pixel-wise Binary Supervision for Face Presentation Attack Detection. URL: https://arxiv.org/pdf/1907.04047v1.pdf.

2. Hong Z., Lin Y., Liu H., Yeh. Y. National Taiwan University. Domain-generalized face anti-spoofing with unknown attacks. URL: https://arxiv.org/pdf/2310.11758v1.pdf.

3. Liu W., Lin C., Yan Y., Beijing Institute Technology. Enhancing Mobile Face Anti-Spoofing: A Robust Framework for Diverse Attack Types under Screen Flash.

4. Srivatsan K., Naseer M., Nandakumar K. FLIP: Cross-domain Face Anti-spoofing with Language Guidance. URL: https://arxiv.org/pdf/2309.16649v1.pdf.