

УДК 621.377.037.3

Ю.А. Харина

## МЕТОДИКА ОЦІНЮВАННЯ СИСТЕМ ВІДЕОАНАЛІТИКИ НА ПРИКЛАДІ БІБЛІОТЕКИ ОБРАЗІВ I-LIDS<sup>1</sup>

У статті описується методика оцінювання систем відеоаналітики (далі – ВА) за допомогою бібліотеки образів на прикладі британської бібліотеки i-LIDS. Порушується проблематика оцінювання таких систем, описується структура бібліотеки i-LIDS, її склад, а також наводяться сценарії, за якими, власне, оцінюються системи ВА, що тестуються. Розглядаються п'ять рівнів сертифікації систем залежно від їх проходження чи непроходження тих чи інших сценаріїв, а також три різні набори даних у складі бібліотеки для тренування та тестування систем. Наводяться ілюстровані приклади помилок спрацювання систем.

**Ключові слова:** відеоаналітика, i-LIDS, бібліотека образів, відхилення спрацювання.

В статье описывается методика оценивания систем видеоналитики с помощью библиотеки образов на примере британской библиотеки образов i-LIDS. Поднимается проблематика оценивания таких систем, описывается структура библиотеки i-LIDS, ее состав, а также приводятся сценарии, по которым, собственно, и оцениваются тестируемые системы ВА. Рассматриваются пять уровней сертификации систем, в зависимости от прохождения или непрохождения тех или иных сценариев, а также существование в составе библиотеки трех разных наборов данных для целей тренировки и тестирования систем. Приводятся иллюстрированные примеры ошибок срабатывания систем.

**Ключевые слова:** видеоналитика, i-LIDS, библиотека образов, отклонение срабатывания.

This paper discusses a method for evaluation of videoanalytis (VA) systems using British image library i-LIDS. The issues of evaluation of such systems are raised, the libraries' structure and composition are described, and scenarios for which evaluated VA-systems have been tested are stated. Five levels of certification systems, depending on their passage or failure of certain scenarios, as well as the existence in the library of three different sets of data for training and testing systems are considered. Illustrated examples of recall biases of systems are given.

**Keywords:** videoanalytics, i-LIDS, image library, recall bias.

### 3. Оцінка системи

#### 3.1. Випробування системи відеоаналітики Міністерством внутрішніх справ ВБ

Науковий відділ Міністерства внутрішніх справ Великої Британії консультує Уряд ВБ про ефективність різних систем відеоспостереження на основі результатів звичайних сценаріїв, заснованих на випробуваннях i-LIDS від імені CPNI.

Системи мають десять секунд для звітування про стан тривоги після того, як подія тривоги почалася на оціночному відеоматеріалі. У цей час множина звітів про тривогу буде знехтувана; подія тривоги буде виявлена або не виявлена.

<sup>1</sup> Закінчення. Початок у попередньому номері.

Після цього десятисекундного вікна будь-які подальші відзвітовані тривоги будуть вважатися “хибними позитивними”.

**Прим.: системи не повинні продовжувати подавати тривогу після закінчення подій тривоги.**

Якщо система видає декілька подій тривоги підряд, тоді за кожні п'ять секунд буде записано тільки одна хибна позитивна тривога.

Якщо на відеоматеріалі виникає разом дві або більше подій тривоги, система повинна створювати окрему тривогу для кожного. З метою випробування системи нема необхідності показувати тип події тривоги, який був виявлений у звіті про тривогу, хоча для робочого застосування це може бути бажано.

### 3.2. Відстеження об'єктів

Вихідна оточуюча рамка слідування системи буде порівнюватися з відміченими об'єктами польових даних для кожного кадру; це порівняння повинно виконуватися в формі порівняння центроїд-центроїд. Відстань до центроїду відстежувальної оточуючої рамки при порівнянні з оточуючою рамкою польових даних буде визначати, чи є слід рамки для цього кадру тривогою дійсною позитивною (ДП), хибною позитивною (ХП) чи хибною негативною (ХН).

Під час підрахунку відстані між поверненим центроїдом оточуючої рамки та центроїдом оточуючої рамки польових даних повинен використовуватися Євклідов підрахунок відстані; це виражається такою формулою:

$$\sqrt{(X_{MCTS} - X_{GT})^2 + (Y_{MCTS} - Y_{GT})^2}$$

де  $X_{MCTS}$  та  $Y_{MCTS}$  – координати центроїду рамки в кадрі, що належать до оцінювальних даних.

$X_{GT}$  та  $Y_{GT}$  – координати центроїду рамки в кадрі, що належать до польових даних.

Результуюча відстань повинна бути між відстанню, що визначена як половина найкоротшої сторони оточуючої рамки польових даних для кадру, що визначений як дійсний позитивний (true positive, TP). Якщо відстань між центроїдами більша, аніж відстань, що вимагається, кадр визначається як хибний позитивний (false positive, FP) або хибний негативний (false negative, FN). Приклади кожного результату показані на рис. 1–4.



Рис. 1. Приклад дійсного позитивного результату, де оцінювальний центроїд рамки знаходиться в межах відстані  $x$  до центроїду рамки польових даних

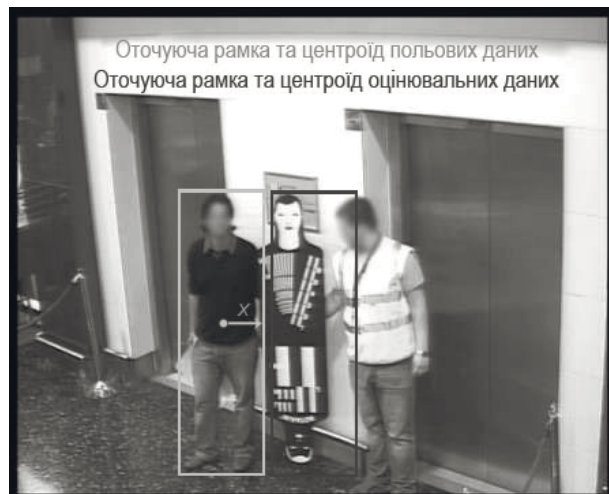


Рис. 2. Приклад хибного позитивного та хибного негативного результатів, де центроїд оцінюваної оточуючої рамки виходить за межі відстані  $x$  до центроїду оточуючої рамки польових даних

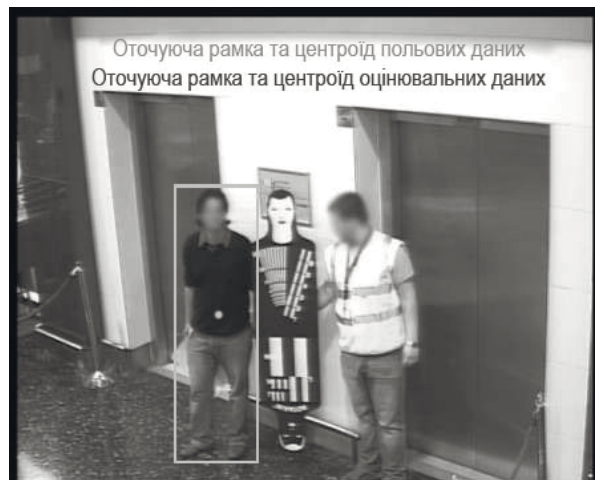


Рис. 3. Приклад хибного негативного результату, де система, що проходить випробування, не повернула результат оточуючої рамки, коли були наявні відповідні польові дані

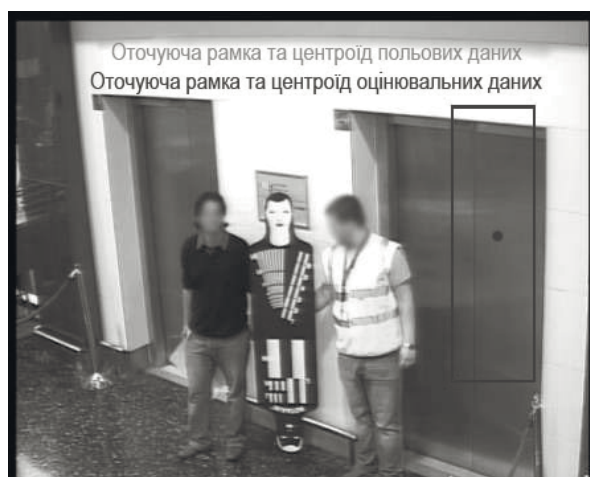


Рис. 4. Приклад хибного позитивного результату, де випробувана система повернула результат оточуючої рамки, що не відповідає наявним польовим даним

Кожний дійсний позитивний, хибний позитивний та хибний негативний результат для кадру був урахований та доданий до остаточного виміру  $F_1$ . Саме за цим виміром, що становить зважене середнє гармонічне системних значень “спрацювання” та “точності”, і оцінюється продукційність систем відеоаналітики за кожним сценарієм [2]:

$$FinalF_1 = \frac{2(Recall \times Precision)}{Recall + Precision},$$

де в остаточному результаті  $F_1$ :

$$Recall = \frac{TotalTP}{TotalTP + TotalFN}$$

$$Precision = \frac{TotalTP}{TotalTP + TotalFP},$$

де  $TotalTP$  – загальна кількість дійсних позитивних спрацювань;

$TotalFN$  – загальна кількість хибних негативних спрацювань;

$TotalFP$  – загальна кількість хибних позитивних спрацювань.

Цей спосіб буде створювати метрику в цілому, схоже на сценарій виявлення подій I-LIDS, який може бути використаний для визначення якості алгоритму на всьому наборі даних.

Система вимагає безперервно відстежувати задану ціль, і якщо безперервність загублена, то для періоду довшого, аніж чотири секунди, в одному наданому прикладі система буде це розглядати як збій тієї конкретної доріжки.

Ураховуючи положення угоди та умов, що регулюють процес оцінки, системи відеоаналітики, що демонструють виміри продукційності  $F_1$  понад встановлені межі, будуть рекомендовані для практичного використання у відповідних сценарії та ролі. Ці системи відеоспостереження будуть вписані в каталог допущеного охоронного обладнання, яке використовується Урядом та іншими частинами критичної національної інфраструктури ВБ.

Значення  $F_1$ , яке повинно бути отримано для того, щоб кваліфікувати системи для практичних рекомендацій, не опубліковується.

### 3.3. Виявлення подій

Системи відеоспостереження можуть бути оцінені за допомогою CAST для ролей “Операційної тривоги” або “Запис події”. У минулому системи забезпечували виявлення в реальному часі підозрілих подій, які повинні бути розглянуті людиною-контролером. Згодом вони діяли як тригер для запису підозрілих подій, де всі отримані записи повинні бути проаналізовані пізніше.

Як було зазначено вище, CAST оцінює продуктивність систем, засновану на вимірі  $F_1$ . Він залежить від параметру, який називається відхилення спрацювання ( $\delta$ ), що визначає вплив на значення  $F_1$  частоти помилкової тривоги відносно частоти виявлення (спрацювання).

Система повинна відповідати оціночному критерію приймання до експлуатації за показником  $F_1$  в 0,7 на стадії розробки. Автори систем, що оцінюються, повинні бути проінформовані про те, що системи мають відповідати значно більш жорстким рівням продуктивності, щоб досягнути рівень продуктивності i-LIDS.

Кожна система відеоаналітики, що представлена з повним набором даних оціночного відеоматеріалу, створює певну кількість:

- дійсних позитивних тривог (a)  
(тривоги системи у відповідь на унікальну подію тривоги);
- хибних позитивних тривог (b)  
(тривоги системи без наявності унікальної події тривоги);
- хибних негативних тривог (c)  
(унікальні події тривоги не призводять до тривоги системи).

Спрацювання (швидкість виявлення) виражається формулою:  $r = \alpha / (\alpha + c)$ . Точність (ймовірність тривоги, яка буде унікальною) виражається формулою:  $p = \alpha / (\alpha + b)$ .

Отже, показник  $F_1$  можна отримати таким чином:

$$F_1 = \frac{(\alpha + 1)rp}{r + \alpha p},$$

де  $a$  – відхилення спрацювання (“recall bias”); зваженість відхилення відносно точності оголошується в кожному визначенні сценарію [1].

Більш високе значення відхилення спрацювання використовується для доступу до систем для ролі “Запис подій”, оскільки при цій ролі хибні спрацювання є менш значущою проблемою. Знання значень відхилення спрацювання дозволяє виробникам оптимізувати свої системи для всіх ролей при випробуванні CAST.

Ураховуючи положення угоди та умов, що регулюють процес оцінки, системи відеоаналітики, що демонструють виміри продуктивності  $F_1$  понад встановлені межі, будуть рекомендовані для практичного використання у відповідних сценаріях та ролі. Ці системи відеоспостереження будуть вписані в каталог допущеного охоронного обладнання, яке використовується Урядом та іншими частинами критичної національної інфраструктури ВБ.

Отже, фактично оцінювання системи відеоаналітики є частиною процесу її розвитку. При цьому визначається не тільки її відповідність технічному завданню, але й конкретизується план її розвитку на майбутнє.

При застосуванні у МВС та Національній поліції України подібна методика оцінювання систем відеоаналітики могла б внести системність у процес впровадження систем відеоаналітики. Крім того, оцінювання систем відеоаналітики за допомогою бібліотеки образів – це найбільш близький до реальних умов спосіб оцінки, і при цьому такий, що ставить системи, які оцінюються, в повністю рівні умови. Ураховуючи досвід країн, що мають впроваджені системи оцінювання типу i-LIDS, в Україні також необхідно впроваджувати аналогічні методики оцінювання.

#### СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ

1. Imagery Library for Intelligent Detection Systems – i-LIDS user guide / 10/11, v. 4.9.
2. C.J. van Rijsbergen. Information Retrieval. Butterworths, London, 1979.

Отримано 31.03.2017

Рецензент Марченко О.С., к.т.н.