

# СИСТЕМА МНОГОКАМЕРНОЙ ТЕМПОРАЛЬНОЙ ИДЕНТИФИКАЦИИ И ТРЕКИНГА ЧЕЛОВЕКА НА ПРЕДПРИЯТИИ

Леус А. В.<sup>1</sup>, Татарина Е. А.<sup>2</sup>, Ефремов В. А.<sup>3</sup>,  
Зуев В. А.<sup>4</sup>, Мангазеев Д. И.<sup>5</sup>, Холодняк И. В.<sup>6</sup>

**Ключевые слова:** переработка визуальной информации, камера, темпоральная идентификация, трекинг, алгоритмы детектирования и локализации, отслеживание объектов, нейронная сеть, макет, эффективность.

## Аннотация

**Цель работы:** разработка макета системы темпоральной идентификации человека и отслеживания временно идентифицированного человека на предприятии по установленным с разными ракурсами камерами.

**Методы:** формально-математический аппарат решения задачи оптимизации архитектуры детектора для локализации объектов с высокой вероятностью; алгоритм слежения, основанный на модифицированном фильтре Калмана, повышающий устойчивость идентификации и эффективного формирования траекторий движения объектов.

**Результаты:** решена задача детектирования фигур людей в видеопотоке, полученном с нескольких камер, в режиме времени, близком к реальному, и идентификации их в рамках определенного временного интервала; разработан формально-математический аппарат решения задачи оптимизации параметров нейронной сети для обеспечения максимально близкого к истинному предсказания; выполнено обучение моделей и оценка их точности; проведено тестирование макета системы многокамерной темпоральной идентификации и трекинга человека на предприятии, показавшее, что разработанный макет обеспечивает идентификацию людей в пределах заданного времени с высокой точностью.

DOI: 10.21681/1994-1404-2022-3-31-39

## Введение

Отслеживание движения объектов в видеопотоке становится необходимым условием выполнения различных практических задач. Контроль движения в контексте помещения и отслеживание движения различных объектов в видеопотоке [7, 9] требуется в системах безопасности и охранных системах, маркетинговых задачах, системах автоматизации производства и др. В каких-то областях целью является слежение за положением объекта в реальном времени в определенном пространстве, в других — необходимо

осуществлять непрерывный анализ траекторий объектов с целью извлечения более сложной комплексной информации [4]. Решение данных задач становится возможным с помощью применения оптико-электронных систем технического зрения, способных без участия человека-оператора выделять во входном видеопотоке объекты интереса и осуществлять слежение за данными объектами [10]. Обработка информации, получаемой оптико-электронной системой, осуществляется в автоматическом режиме. Оптико-электронная система приобретает роль «органов зрения», ключевым моментом построения которых является необходимость реализации качественных алгоритмов обнаружения и слежения за объектами [1, 2, 5].

<sup>1</sup> **Леус Андрей Владимирович**, кандидат технических наук, ведущий программист лаборатории цифровых систем специального назначения Московского физико-технического института (МФТИ), г. Москва, Российская Федерация.

E-mail: leus.av@mipt.ru

<sup>2</sup> **Татарина Елена Александровна**, кандидат технических наук, аналитик лаборатории цифровых систем специального назначения МФТИ, г. Москва, Российская Федерация.

E-mail: tatarinova.ea@mipt.ru

<sup>3</sup> **Ефремов Владислав Александрович**, программист лаборатории цифровых систем специального назначения, аспирант МФТИ, г. Москва, Российская Федерация.

E-mail: efremov.va@mipt.ru

<sup>4</sup> **Зуев Виктор Александрович**, студент 5 курса МФТИ, г. Москва, Российская Федерация.

E-mail: zuev.va@phystech.edu

<sup>5</sup> **Мангазеев Даниил Игоревич**, студент 4 курса МФТИ, г. Москва, Российская Федерация.

E-mail: mangazeev.di@phystech.edu

<sup>6</sup> **Холодняк Иван Витальевич**, студент 5 курса МФТИ, г. Москва, Российская Федерация.

E-mail: kholodnyak.iv@phystech.edu

Задача отслеживания объектов<sup>7</sup> или целей<sup>8</sup> играет важную роль при создании *систем компьютерного зрения*. Целями для отслеживания могут быть различные живые объекты, например, пешеходы на улице или люди, передвигающиеся в ограниченном пространстве, транспортные средства на дороге, спортсмены, группы животных.

Задачу выделения и сопровождения объектов на последовательности видеок кадров можно определить как задачу оценивания траектории объекта в плоскости изображения в процессе движения объекта внутри сцены<sup>9</sup>. Другими словами, алгоритм слежения должен каждому сопровождаемому объекту присвоить метку, не изменяющуюся на протяжении всей последовательности видеок кадров. Дополнительно, в зависимости от конкретной прикладной области, может требоваться определение ориентации, формы или площади объекта. С целью увеличения скорости обработки данных при программных и аппаратных реализациях алгоритмов такого типа вводят ряд упрощений [6]. Одно из самых частых — введение ограничений на тип движения и вид объектов. Например, практически все алгоритмы слежения предполагают, что траектория движения объекта достаточно гладкая, без резких изменений (разрывов первого и второго рода), и, более того, движение объекта ограничивают движением с постоянной скоростью или ускорением. Для упрощения задачи также используют априорную информацию о количестве и размере объектов или внешнем виде и форме объектов слежения.

В настоящее время существует достаточно много различных групп алгоритмов детектирования и локализации: блочные методы, методы оптического потока, фазовой корреляции, глобальной оценки движения, слежения за особенностями, многокадровой оценки движения, которые различаются по следующим основным характеристикам: способу представления объекта слежения, набору признаков изображения, используемому для слежения; используемому методу выделения объектов, способу определения траектории движения объекта. Одним из наиболее перспективных методов детектирования и локализации объектов на изображениях является использование для этих целей искусственных *нейронных сетей* [3, 11].

Целью исследования является разработка макета системы темпоральной идентификации человека (в пределах заранее заданного интервала времени) и отслеживания временно идентифицированного человека на предприятии по установленным с разными

ракурсами камерами (многокамерная темпоральная идентификация и трекинг человека на предприятии).

### Постановка задачи

Для достижения поставленной цели разработана архитектура и модель, а также выполнена проверка работоспособности детектора, обеспечивающего достаточную вероятность локализации объектов на изображении с учетом требований быстродействия системы.

Основным назначением макета системы многокамерной темпоральной идентификации и трекинга человека на предприятии является обеспечение возможности идентификации человека в пределах заданного интервала времени, например, одного рабочего дня с учетом достижимых характеристик, а также возможности отслеживания идентифицированного человека в течение заданного интервала времени на предприятии по установленным с разными ракурсами камерам.

Состав разрабатываемого макета системы многокамерной темпоральной идентификации и трекинга человека на предприятии включает следующие ключевые элементы:

- *видеокамеры* для осуществления видеомониторинга;
- *вычислитель* для обработки изображений, получаемых с видеокамер, посредством разработанных алгоритмов (в качестве вычислителя используется персональный компьютер с видеокартой типа RTX, для реализации алгоритмов — достаточно производительной бытовой видеокартой RTX 3080 или ее аналога);
- *алгоритм темпоральной идентификации человека*, обеспечивающий идентификацию каждой обнаруженной человеческой фигуры и присвоение ей идентификационного номера;
- *алгоритм отслеживания временно идентифицированного человека*, обеспечивающий при каждом появлении указанного объекта обозначение его собственным темпоральным идентификатором.

Каждая видеокамера осуществляет запись видео. Поля зрения видеокамер могут пересекаться. Данные с камер передаются на вычислитель, где осуществляется их логическая обработка с помощью программного обеспечения, основанного на нейронных сетях. На каждом кадре с помощью заранее обученного на обнаружение и выделение людей ограничивающими рамками детектора выделяются человеческие фигуры. Каждой человеческой фигуре при первом ее появлении в поле зрения камер в пределах обозначенного интервала времени (например, один час или одни сутки) сопоставляется *темпоральный идентификационный номер*. Поскольку поля зрения камер в макете специальным образом пересекаются, то в рамках модельной задачи фигура при первом появлении всегда определяется минимум на двух камерах с разными ракурсами. После фиксирования темпорального идентификацион-

<sup>7</sup> Отслеживание объектов (англ. object tracking) — задача прогнозирования положения или другой важной информации о движущихся объектах в видеоизображении.

<sup>8</sup> Luo W., Xing J, Milan A., Zhang X., Liu W., Kim T.-K. Multiple object tracking: a literature review. URL: <https://arxiv.org/pdf/1409.7618.pdf> (дата обращения: 15.08.22).

<sup>9</sup> Ristani E., Tomasi C. Features for multi-target multi-camera tracking and re-identification. URL: <https://arxiv.org/pdf/1803.10859.pdf> (дата обращения: 15.08.22).

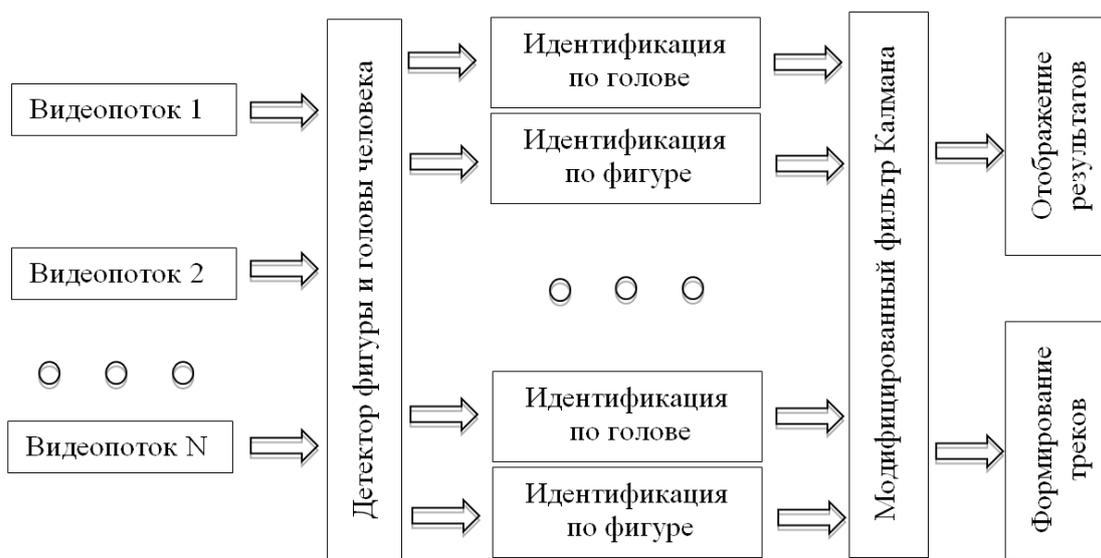


Рис. 1. Схема работы макета системы многокамерной темпоральной идентификации и трекинга человека на предприятии

ного номера для данного объекта в течение рассматриваемого промежутка времени, при каждом появлении указанного объекта в поле зрения любой из трех установленных камер объект обозначается своим темпоральным идентификатором. При первом появлении в поле зрения двух камер нового человека ему присваивается другой темпоральный идентификатор. Далее для каждого изображения, где будет найден один или несколько объектов, все объекты выделяются ограничивающими рамками, а каждой рамке соответствует верный темпоральный идентификатор.

В результате работы система обеспечивает обнаружение и выделение попадающих в поле зрения человеческих фигур ограничивающими рамками с указанием темпорального идентификационного номера. Темпоральные идентификаторы объектов действуют в пределах заранее определенного промежутка времени, например, один рабочий день.

### Описание работы системы

Одной из основных задач исследования является задача отслеживания объектов, состоящая из двух подзадач: последовательной детекции от кадра к кадру и идентификации. В основе *первой* подзадачи лежит обнаружение объекта на кадре, *второй* — определение уникального идентификатора, который необходимо сопоставить с идентификаторами предыдущих состояний. Общая схема работы системы представлена на рис. 1.

Для первичного запоминания человека осуществляется первоначальная опорная запись, на которой видео с объектом записывается таким образом, что его видно одновременно с нескольких сторон. На основании данной записи с помощью нейросетевых алгоритмов формируются векторы признакового представ-

ления данного объекта, которые записываются в базу данных, объекту присваивается идентификатор. Поскольку запись на данном этапе ведется одновременно с нескольких сторон, то в рамках полученных векторов будут фиксироваться все визуально доступные отличительные признаки объекта.

При последующей работе системы вычислитель обрабатывает видеопотоки одновременно с нескольких камер. На каждом кадре нейросеть детектирует фигуру человека, с помощью нейросетевого алгоритма формирует вектор признакового представления данного объекта, сравнивает данный вектор с имеющимися в базе данных. Если полученный вектор достаточно близок к одному из имеющихся векторов в пределах заранее определенного порога, то найденному объекту присваивается соответствующий идентификатор из базы.

Для повышения точности распознавания объекта при детектировании фигуры человека нейросеть одновременно детектирует голову человека. Процедура с головой в точности повторяет процедуру с телом человека. Таким образом, непрерывная идентификация человека проводится независимо с помощью двух идентифицирующих нейронных сетей, отдельно обученных на головы и фигуры людей. Это позволяет лучше различать людей в одинаковой одежде.

Для повышения *устойчивости* идентификации и эффективного формирования траекторий движения объектов используется алгоритм слежения, основанный на модифицированном фильтре Калмана. Фильтр Калмана достаточно активно используется для трекинга объектов<sup>10</sup> и дает оценку состояния линейной систе-

<sup>10</sup> Beymer D., Konolige K., Park M. Real-time tracking of multiple people using continuous detection. IEEE Frame Rate Workshop. 1999. No 6. P. 1–8; Broida T. J., Chellappa R. Estimation of Object Motion Parameters from Noisy Images. IEEE Trans. Pattern Anal. Mach. 1986. Vol. 8, No 1. P. 90–99.

мы в предположении, что сами состояния распределены по нормальному закону. Процедура состоит из двух шагов:

*первый шаг* — прогноз или экстраполяция, используется для предсказания новых состояний переменных:

$$\bar{X}^t = \bar{\Sigma}^t = D, \\ \bar{\Sigma}^{t-1}D^T + Q^tDX^{t-1} + W,$$

где  $\bar{X}^t$  и  $\bar{\Sigma}^t$  — соответственно состояние переменных и их дисперсия в момент времени  $t$ ;  $D$  — матрица переходов, которая определяет возможность изменения состояния при переходе от момента времени  $t - 1$  к  $t$ ;  $Q$  — ковариационная матрица шума  $W$ ;

*второй шаг* — коррекция, использующая текущие наблюдения  $Z^t$  для оценки отклонения наблюдения, полученного на шаге  $t$ , от ожидаемого при произведенной экстраполяции:

$$K^t = \bar{\Sigma}^t M^t [M^t \bar{\Sigma}^t M^t + R^t]^{-1}, \\ X^t = \bar{X}^t + K^t [Z^t - M^t \bar{X}^t], \\ \Sigma^t = \bar{\Sigma}^t - K^t M^t \bar{\Sigma}^t,$$

где  $M$  — матрица наблюдений;  $K$  — оптимальная по Калману матрица коэффициентов усиления, формирующаяся на основании ковариационных матриц, имеющейся экстраполяции вектора состояния и полученных измерений (посредством ковариационной матрицы вектора отклонения).

Для выбора оптимальной архитектуры детектора для локализации объектов с высокой вероятностью, проведены исследования алгоритмов машинного обучения из серии YOLOv4-v5<sup>11</sup> (You Only Look Once) [12]. YOLOv5 — это последняя версия семейства алгоритмов машинного обучения, основанных на обнаружении объектов, что означает, что детектор ищет на изображении определенные объекты. YOLOv4-v5 относятся к одностадийным детекторам, которые используют сверхточные нейронные сети для обнаружения объектов. Время инференса на одно изображение составляет 12,5 мс на видеокarte уровня NVIDIA GeForce RTX 3080. То есть для обработки 30 кадров требуется около 0,3 с времени, что делает возможным внедрение таких алгоритмов в задачи отслеживания объектов.

### Формально-математический аппарат решения задачи оптимизации

Задача оптимизации представляет собой задачу подбора таких параметров нейронной сети, при которых по данным, которые подаются на вход модели, она могла бы приблизить предсказания максимально точно к истинным значениям. Задача оптимизации формулируется следующим образом.

Существуют истинные ответы  $y$  и предсказание модели  $f(x, p)$ . Задан функционал потерь:

$$F_{loss} = \sum_{i=1}^N L(y_i, f(x_i, w)),$$

где  $N$  — размер выборки;  $f$  — функция, соответствующая архитектуре нейросети;  $w$  — набор параметров нейросети.

Функционал потерь принимает минимальное значение при следующем условии:

$$w_{opt} = \arg \min \{ \lim_w F_{loss} \}.$$

В качестве функции потерь могут использоваться следующие основные функции [8]:

– кусочно-линейная (*hinge loss*):  
 $l(y, \hat{y}) = \max(0, 1 - y\hat{y}).$

– логистическая:

$$l(y, \hat{y}) = -\hat{y} \ln(y) - (1 - \hat{y}) \ln(1 - y).$$

– функция Хьюбера:

$$l(y, \hat{y}) = \begin{cases} 0,5(\hat{y} - y)^2; & |\hat{y} - y| < \delta \\ \delta(|\hat{y} - y| - 0,5\delta) & \text{иначе} \end{cases}$$

– квадратичная:

$$l(y, \hat{y}) = 0,5(\hat{y} - y)^2.$$

Для обучения нейронной сети используется метод градиентного спуска [8]:

$$w_i = w_{i-1} - \alpha \sum_{i=1}^N \nabla_w L(y_i, f(x_i, w)),$$

где  $\alpha$  — темп обучения.

Для ускорения процесса обучения можно использовать метод упругого распространения (*resilient propagation*)<sup>12</sup>. Данный метод основан на «эпохальном»<sup>13</sup> («вальдорфском») обучении. С целью корректировки весов в нем применяются частные производные. Для определения величины коррекции используется следующее выражение:

$$\Delta_{ij}^{(t)} = \begin{cases} \eta^+ \Delta_{ij}^{(t)}, \frac{\partial E^{(t)}}{\partial w_{ij}} \frac{\partial E^{(t-1)}}{\partial w_{ij}} > 0 \\ \eta^- \Delta_{ij}^{(t)}, \frac{\partial E^{(t)}}{\partial w_{ij}} \frac{\partial E^{(t-1)}}{\partial w_{ij}} < 0 \\ 0 < \eta^- < 1 < \eta^+. \end{cases}$$

Изменение знака частной производной по весу  $w_{ij}$  свидетельствует о слишком большом последнем изме-

<sup>11</sup> Bochkovskiy A., Wang C.Y., Meng B., Liao H.-Y. M. Optimal Speed and Accuracy of Object Detection. URL: <https://arxiv.org/pdf/2004.10934.pdf> (дата обращения: 15.08.22).

<sup>12</sup> Riedmiller M., Braun H. A. Direct adaptive method for faster backpropagation learning: the RPROP algorithm. IEEE International Conference on Neural Networks. 1993. P. 586–591.

<sup>13</sup> Морковин А.М. Зарубежный опыт функционирования вальдорфских школ на современном этапе // Мир науки, культуры, образования. 2018. № 4(71). С. 150—151.

Таблица 1

Сравнительный анализ моделей

Модель	Время инференса изображения, мс	Метрика точности <sup>15</sup> , усл. ед.
Модель 1	10,5	0,828
Модель 2	12,5	0,846
Модель 3	12,5	0,838

Таблица 2

Результаты оценки точности реидентификации отдельно по фигуре человека и по голове человека, обученных на двух разных датасетах

Модель	Истина	Потери	Ошибка
Фигура человека*	0,912	0,025	0,064
Голова человека*	0,586	0,040	0,374
Голова человека**	0,605	0,355	0,040
Голова человека***	0,778	0,180	0,042

**Примечание:** \* — обучение на датасете *market*; \*\* — обучение на датасете *cfp*; \*\*\* — обучение на датасете *cfp\_YouTube*.

нении, следовательно, величину изменения необходимо уменьшить на  $\eta^-$  и вернуть предыдущее значение весового коэффициента:

$$\Delta w_{ij}(t) = \Delta w_{ij}(t) - \Delta_{ij}^{(t-1)}, \quad 0 < \eta^- < 1 < \eta^+.$$

В том случае, если знак частной производной не изменился, необходимо увеличить величину коррекции на  $\eta^+$ . Для того чтобы веса не принимали слишком больших или слишком малых значений, используются ограничения величины коррекции. Фиксирование основных параметров коррекции позволяет избежать настройки глобальных параметров. Коррекция весов осуществляется в соответствии со следующим правилом:

$$\Delta w_{ij}(t) = \begin{cases} -\Delta_{ij}^{(t)}, & \frac{\partial E^{(t)}}{\partial w_{ij}} > 0 \\ +\Delta_{ij}^{(t)}, & \frac{\partial E^{(t)}}{\partial w_{ij}} < 0 \\ 0, & \frac{\partial E^{(t)}}{\partial w_{ij}} = 0 \end{cases}.$$

В случае положительной производной ошибка возрастает, весовой коэффициент уменьшается на величину коррекции, в противном случае коэффициент увеличивается. Подстройка весов осуществляется в соответствии с выражением:

$$w_{ij}(t + 1) = w_{ij}(t) - \Delta w_{ij}(t).$$

### Обучение моделей

Для тренировки и тестирования использовался датасет *CrowdHuman*<sup>14</sup>, содержащий 15 000 фотографий людей. В общей сложности датасет содержит 339 565 уникальных объектов. Для каждого человека выполнена разметка на фигуру и голову.

*YOLOv5* предоставляет четыре разные версии сети для обучения модели: маленький (*YOLOv5s*), средний (*YOLOv5m*), большой (*YOLOv5l*) и очень большой (*YOLOv5x*). Были натренированы 2 модели размера *m* и *s*. В сравнительный анализ вошла также доступная версия *yolov5m\_CrowdHuman*. Результаты сравнительного анализа представлены в табл. 1.

Для создания модели 1 использовалась нейронная сеть *YOLOv5s*, для создания модели 2 — нейронная сеть *YOLOv5m*, для модели 3 — *yolov5m\_CrowdHuman*.

Оценка точности работы системы с учетом детекции и идентификации проводилась с помощью двух модулей.

Первый модуль создает таблицы по видеофайлам в папках *query* и *gallery*, видеофайлы должны содержать изображения только одного конкретного человека. Скрипт создает векторы признакового представления.

Второй модуль создает результирующую таблицу по предыдущим таблицам, в которой содержится процент попадания в истинное значение по каждому классу и

<sup>14</sup>Shao S., Shao Z., Li B., Xiao T., Yu G. Zhang X., Sun J. CrowdHuman: A Benchmark for Detecting Human in a Crowd. URL: <https://arxiv.org/pdf/1805.00123.pdf> (дата обращения: 15.08.22).

<sup>15</sup>mAP[0,5] (mean Average Precision — среднее значение точности) — метрика точности идентификации при обнаружении цели.

Результаты предобучения на фигурах человека

Модель	Истина	Потери	Ошибка
MOT17_sbs_s50	0,900	0,084	0,017
MOT20_sbs_s50	0,917	0,045	0,038

Обучение моделей библиотеки *FastReid* на классе «голова»

Обучающий датасет	Количество уникальных голов, шт.	Количество изображений, шт.	Истина	Потери	Ошибка
MOT17	140	6900	0,846	0,049	0,105
MOT17 + DukeMTMC	400	9900	0,872	0,078	0,051
MOT17 + DukeMTMC + YouTube faces	3200	57400	0,877	0,075	0,048

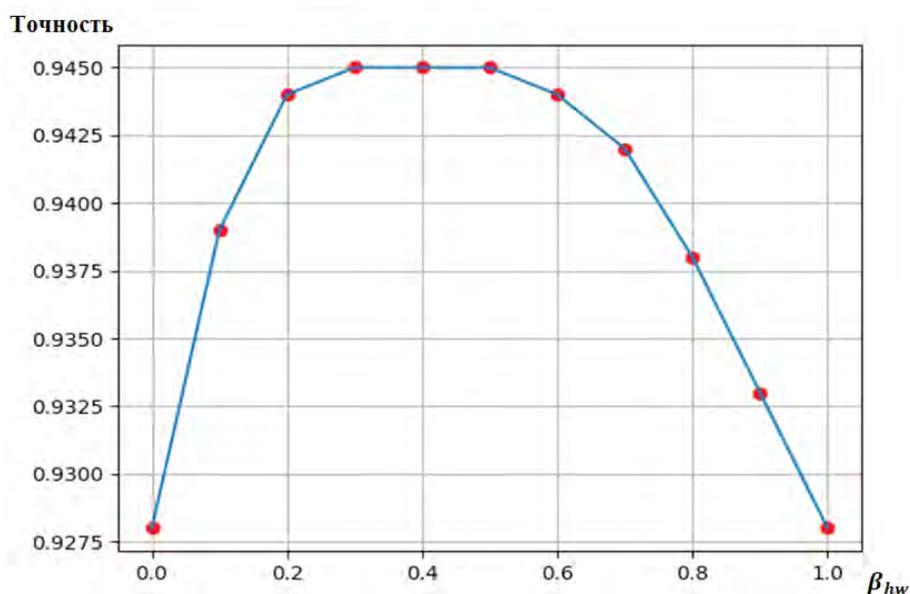


Рис. 2. Результаты исследования зависимости точности от коэффициента предсказаний

среднее по всем классам. Результаты оценки точности представлены в табл. 2.

В табл. 2 «истина» означает долю верно идентифицированных объектов, «потери» — долю неидентифицированных объектов, «ошибка» — долю ошибочно идентифицированных объектов.

Предобучение моделей библиотеки *FastReid*<sup>16</sup> осуществлялось на фигуре человека. *FastReid* имеет модульную структуру, обладает высокой гибкостью и масштабируемостью. Результаты предобучения представлены в табл. 3.

Предобучение осуществлялось на наборе данных для многообъектного слежения *MultipleObjectTracking*

(MOT)<sup>17</sup> версий 17 и 20, позволяющих сравнивать эффективность модели. В качестве модели в табл. 3 использовалась модель *FastReID SBS\_s50*.

Обучение моделей библиотеки *FastReID* осуществлялось на классе «голова». Результаты обучения представлены в табл. 4.

Для создания датасета 1 (см. табл. 4) использовался набор данных *MOT17*, датасет 2 включал наборы данных *MOT17* и *DukeMTMC*<sup>18</sup>, датасет 3, кроме наборов данных *MOT17* и *DukeMTMC*, дополнительно расширен данными *YouTube faces*<sup>19</sup>.

<sup>16</sup> He L., Liao X., Liu W., Liu X., Cheng P. Mei T. FastReID: A PyTorch Toolbox for General Instance Re-identification. URL: <https://arxiv.org/pdf/2006.02631.pdf>

<sup>17</sup> MultipleObjectTracking. URL: <https://motchallenge.net>

<sup>18</sup> DUKE MTMC. URL: [https://exposing.ai/duke\\_mtmc](https://exposing.ai/duke_mtmc)

<sup>19</sup> YouTube faces. URL: <https://www.cs.tau.ac.il/~wolf/ytfaces>

Результаты испытаний макета

Описание датасета	Количество уникальных людей, чел.	Истина	Потери	Ошибка
Корректно записанные видео	30	0,956	0,016	0,028
Собранные видео с YouTube	103	0,554	0,307	0,140

Итоговый результат идентификатора состоит из идентификации по голове  $B$  и идентификации по фигуре человека  $H$ :

$$R = B(1 - \beta_{hw}) + H\beta_{hw},$$

где  $\beta_{hw}$  — коэффициент, определяющий вес каждого из предсказаний.

На рис. 2 представлена кривая зависимости точности идентификации от коэффициента  $\beta_{hw}$ .

Как видно из графика (см. рис. 2), оптимальное значение коэффициента  $\beta_{hw} = 0,4$ , а точность, соответствующая данному параметру, составляет 0,945 (доля верно идентифицированных объектов).

### Экспериментальные исследования

Для проведения тестирования макета системы темпоральной идентификации человека установлены следующие параметры: функция связи — линейная; параметр  $\beta_{hw} = 0,4$ .

Результаты экспериментов в виде усредненных озонамительных данных представлены в табл. 5.

### Заключение

Таким образом, разработан макет системы темпоральной идентификации человека в пределах заранее заданного интервала времени и отслеживания временно идентифицированного человека на предприятии по установленным с разными ракурсами камерам.

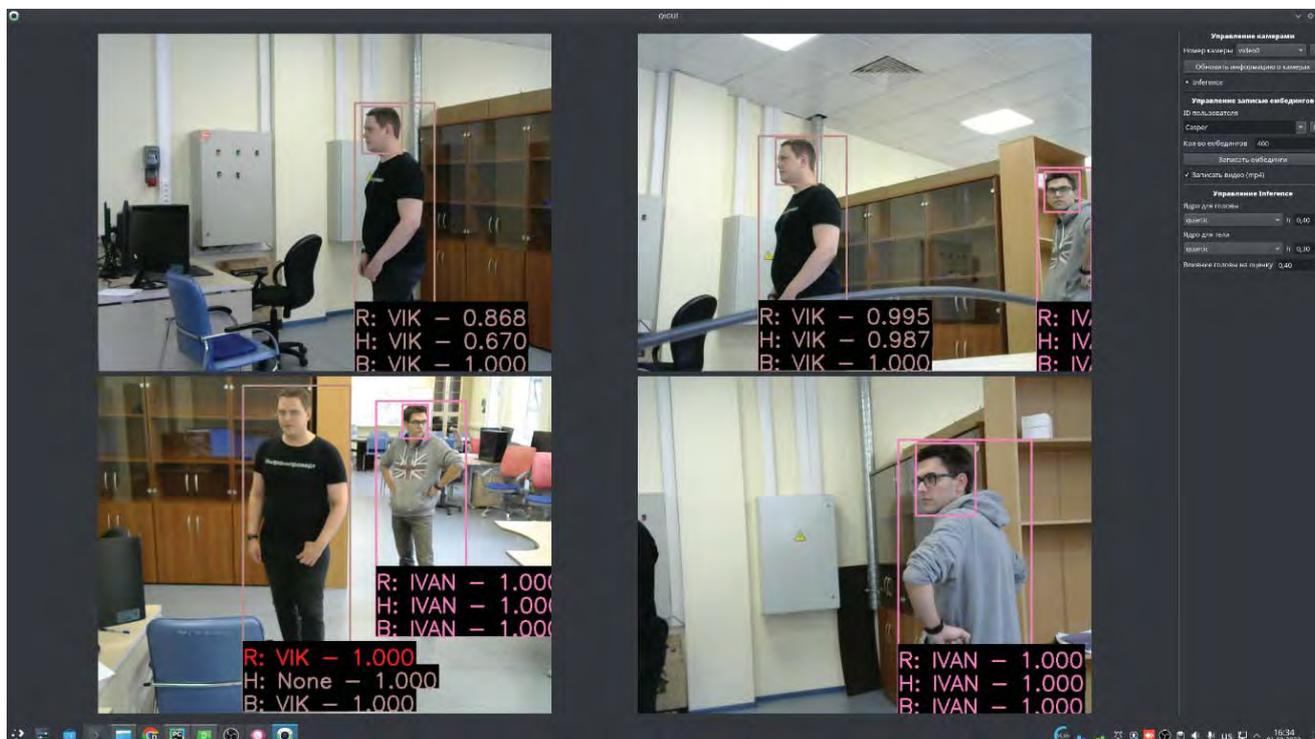


Рис. 3. Пример работы макета системы темпоральной идентификации человека

Выполнены: обоснованный выбор архитектуры и модели, а также проверка работоспособности нейросетевого детектора, обеспечивающего достаточную вероятность локализации объектов на изображении с учетом требований быстродействия системы. Разработан формально-математический аппарат решения задачи оптимизации параметров нейронной сети для обеспечения

максимально близкого к истинному предсказания. Выполнено обучение моделей и оценка их точности.

Проведено тестирование макета системы многокамерной темпоральной идентификации и трекинга человека на предприятии, показавшее, что разработанный макет обеспечивает идентификацию людей в пределах заданного времени с высокой точностью.

*Рецензент: **Алексеев Владимир Витальевич**, доктор технических наук, профессор, член-корреспондент РАН, заведующий кафедрой информационных систем и защиты информации Тамбовского государственного технического университета, г. Тамбов, Российская Федерация.*

*E-mail: [vvalex1961@mail.ru](mailto:vvalex1961@mail.ru)*

### Литература

1. Бурый А.С. Информационно-математическое обеспечение контроля качества компьютерных программ // Правовая информатика. 2019. № 2. С. 15—24. DOI: 10.21681/1994-1404-2019-2-15-25 .
2. Гаврилов Д.А. Интеллектуальные программно-аппаратные решения для автоматизированных оптико-электронных систем реального времени // Правовая информатика. 2021. № 3. С. 14—24. DOI: 10.21681/1994-1404-2021-3-14-24 .
3. Гаврилов Д.А. Нейросетевой алгоритм автоматического обнаружения и сопровождения объекта интереса в видеосигнале // Тр. 16-й Нац. конф. по искусственному интеллекту (24—27 сентября 2018 г.). В 2-х томах / ФИЦ ИУ РАН. Т. 2. М. : РКП, 2018. С. 188—190.
4. Гаврилов Д.А., Ловцов Д.А. Автоматизированная оптико-электронная система наземно-космического мониторинга для систем безопасности реального времени // Вопросы кибербезопасности. 2020. Т. 39. № 5. С. 41—47.
5. Гаврилов Д.А., Ловцов Д.А. Автоматизированная переработка визуальной информации с помощью технологий искусственного интеллекта // Искусственный интеллект и принятие решений. 2020. № 4. С. 33—46.
6. Канушкин С.В. Алгоритмы адаптивного управления робототехническими системами охранного мониторинга // Правовая информатика. 2020. № 2. С. 28—40. DOI: 10.21681/1994-1404-2020-2-28-40 .
7. Ловцов Д.А. Алгоритмизация в крупномасштабных иерархических эргасистемах // Правовая информатика. 2022. № 1. С. 4—14. DOI: 10.21681/1994-1404-2022-1-04-14 .
8. Ловцов Д.А. Информационная теория эргасистем : монография. М. : РГУП, 2020. 314 с. ISBN: 978-5-93916-887-8.
9. Ловцов Д.А., Гаврилов Д.А. Эффективная автоматизированная оптико-электронная система аэрокосмического мониторинга // Правовая информатика. 2019. № 2. С. 29—35.
10. Ловцов Д.А., Гаврилов Д.А. Моделирование оптико-электронных систем дистанционно пилотируемых аппаратов : монография. М. : Технолоджи-3000, 2019. 164 с.
11. Ловцов Д.А., Гаврилов Д.А., Татаринова Е.А. Эффективная автоматизированная переработка визуальной информации в оптико-электронной системе наземно-космического мониторинга // Профессорский журнал. Сер. Технические науки. 2019. № 3. С. 26—40.
12. Wang Z., Wu Y., Yang L., Thirunavukarasu A., Evison C., Zhao Y. Fast personal protective equipment detection for real construction sites using deep learning approaches. *Sensors*, 2021, vol. 21. URL: <https://doi.org/10.3390/s21103478>

# A SYSTEM FOR MULTI-CAMERA TEMPORAL IDENTIFICATION AND TRACKING OF A PERSON AT THE ENTERPRISE

*Andrei Leus<sup>20</sup>, Elena Tatarinova<sup>21</sup>, Vladislav Efremov<sup>22</sup>, Viktor Zuev<sup>23</sup>, Daniil Mangazeev<sup>24</sup>, Ivan Kholodniak<sup>25</sup>*

<sup>20</sup> **Andrei Leus**, Ph.D. (Technology), Leading Programmer at the Laboratory of Special Purpose Digital Systems of the Moscow Institute of Physics and Technology (MIPT), Moscow, Russian Federation. E-mail: [leus.av@mipt.ru](mailto:leus.av@mipt.ru)

<sup>21</sup> **Elena Tatarinova**, Ph.D. (Technology), Analyst at the Laboratory of Special Purpose Digital Systems of the Moscow Institute of Physics and Technology (MIPT), Moscow, Russian Federation. E-mail: [tatarinova.ea@mipt.ru](mailto:tatarinova.ea@mipt.ru)

<sup>22</sup> **Vladislav Efremov**, Programmer at the Laboratory of Special Purpose Digital Systems of the MIPT, Ph.D. student at the MIPT, Moscow, Russian Federation. E-mail: [efremov.va@mipt.ru](mailto:efremov.va@mipt.ru)

<sup>23</sup> **Viktor Zuev**, 5<sup>th</sup> year student at the MIPT, Moscow, Russian Federation. E-mail: [zuev.va@phystech.edu](mailto:zuev.va@phystech.edu)

<sup>24</sup> **Daniil Mangazeev**, 4<sup>th</sup> year student at the MIPT, Moscow, Russian Federation. E-mail: [mangazeev.di@phystech.edu](mailto:mangazeev.di@phystech.edu)

<sup>25</sup> **Ivan Kholodniak**, 5<sup>th</sup> year student at the MIPT, Moscow, Russian Federation. E-mail: [kholodnyak.iv@phystech.edu](mailto:kholodnyak.iv@phystech.edu)

**Keywords:** visual information processing, camera, temporal identification, tracking, detection and localisation algorithms, object tracking, neural network, model, efficiency.

### Abstract

*Purpose of the work: developing a model of a system for temporal identification of a person and tracking of the temporarily identified person at the enterprise using cameras installed at different angles.*

*Methods used: formal mathematical apparatus for solving the problem of optimising the architecture of a detector for localising objects with a high probability, a tracking algorithm based on a modified Kalman filter which increases the stability of identification and efficient construction of object movement trajectories.*

*Findings: the problem of detecting human figures in a video stream received from several cameras in near real time and identifying them within a certain time interval was solved. A formal mathematical apparatus was developed for solving the problem of optimising the parameters of a neural network for ensuring a prediction being as close as possible to the true one. Training of models and an assessment of their accuracy were carried out. A testing of the model of a system for multi-camera temporal identification and tracking of a person at the enterprise was carried out which showed that the developed model provides a highly accurate identification of persons within a specified time period.*

### References

1. Buryi A.S. Informatsionno-matematicheskoe obespechenie kontrolya kachestva komp'yuternykh programm. Pravovaia informatika, 2019, No. 2, pp. 15–24. DOI: 10.21681/1994-1404-2019-2-15-25 .
2. Gavrilov D.A. Intellektual'nye programmno-apparatnye resheniia dlia avtomatizirovannykh optiko-elektronnykh sistem real'nogo vremeni. Pravovaia informatika, 2021, No. 3, pp. 14–24. DOI: 10.21681/1994-1404-2021-3-14-24 .
3. Gavrilov D.A. Neurosetevoi algoritm avtomaticheskogo obnaruzheniia i soprovozhdeniia ob'ekta interesa v videosig-nale. Tr. 16-i Nats. konf. po iskusstvennomu intellektu (24–27 sentiabria 2018 g.). V 2-kh tomakh. FITs IU RAN, t. 2. M. : RKP, 2018, pp. 188–190.
4. Gavrilov D.A., Lovtsov D.A. Avtomatizirovannaia optiko-elektronnaia sistema nazemno-kosmicheskogo monitoringa dlia sistem bezopasnosti real'nogo vremeni. Voprosy kiberbezopasnosti, 2020, t. 39, No. 5, pp. 41–47.
5. Gavrilov D.A., Lovtsov D.A. Avtomatizirovannaia pererabotka vizual'noi informatsii s pomoshch'iu tekhnologii iskusstvennogo intellekta. Iskusstvennyi intellekt i priniatie reshenii, 2020, No. 4, pp. 33–46.
6. Kanushkin S.V. Algoritmy adaptivnogo upravleniia robototekhnicheskimi sistemami okhrannogo monitoringa. Pravovaia informatika, 2020, No. 2, pp. 28–40. DOI: 10.21681/1994-1404-2020-2-28-40 .
7. Lovtsov D.A. Algoritmizatsiia v krupnomasshtabnykh ierarkhicheskikh ergasistemakh. Pravovaia informatika, 2022, No. 1, pp. 4–14. DOI: 10.21681/1994-1404-2022-1-04-14 .
8. Lovtsov D.A. Informatsionnaia teoriia ergasistem : monografiia. M. : RGUP, 2020. 314 pp. ISBN: 978-5-93916-887-8.
9. Lovtsov D.A., Gavrilov D.A. Effektivnaia avtomatizirovannaia optiko-elektronnaia sistema aerokosmicheskogo monitoringa. Pravovaia informatika, 2019, No. 2, pp. 29–35.
10. Lovtsov D.A., Gavrilov D.A. Modelirovanie optiko-elektronnykh sistem distantsionno pilotiruemykh apparatov : monografiia. M. : Tekhnolodzhi-3000, 2019. 164 pp.
11. Lovtsov D.A., Gavrilov D.A., Tatarinova E.A. Effektivnaia avtomatizirovannaia pererabotka vizual'noi informatsii v optiko-elektronnoi sisteme nazemno-kosmicheskogo monitoringa. Professorskii zhurnal, ser. Tekhnicheskie nauki, 2019, No. 3, pp. 26–40.
12. Wang Z., Wu Y., Yang L., Thirunavukarasu A., Evison C., Zhao Y. Fast personal protective equipment detection for real construction sites using deep learning approaches. Sensors, 2021, vol. 21. URL: <https://doi.org/10.3390/s21103478>