

## **ЭКСПЕРИМЕНТАЛЬНЫЙ ВЫБОР ГИПЕРПАРАМЕТРОВ СВЕРТОЧНОЙ НЕЙРОННОЙ СЕТИ DENSENET-121 ДЛЯ РЕИДЕНТИФИКАЦИИ ЛЮДЕЙ ПО ИЗОБРАЖЕНИЯМ**

**С. А. Игнатьева, Р. П. Богуш**

Полоцкий государственный университет имени Евфросинии Полоцкой,  
Новополоцк, Беларусь, s.ignatieva@psu.by

*Hyperparameter choice is one of the key methods to improve accuracy for people re-identification. For this task, we used a two-stage approach. Our methodology allowed us to find among many possible values the most effective combination for batch size, learning rate and number of epochs of CNNs training for DenseNet-121. Training and testing were performed on the four largest and most used datasets such as Market-1501, DukeMTMC-ReID, PolReID1077 and MSMT17.*

Развитие аппаратной базы систем видеонаблюдения, увеличение пропускной способности каналов связи, облачных решений, методов искусственного интеллекта для обработки информации приводит тому, что видеоаналитика данных все шире используется в распределенных системах видеонаблюдения. Интеллектуальность видеоаналитики заключается в способности автоматически анализировать видеопотоки с целью выявления заданных объектов или оценки их действий. К задаче видеоаналитики относится реидентификация людей. Реидентификация предполагает определение присутствия заданного человека по его изображениям на видеоданных в другом месте или в разное время в пространственно-распределенных системах видеонаблюдения [1]. Однако высокая сложность реидентификации обусловлена рядом мешающих факторов, включая, сложный фон, перекрытия, вариативность поз, неоднозначный внешний вид, пересечение множества траекторий движения и др. (см.рис.1). Поэтому для эффективного решения таких задач актуальна проблема выбора дескриптора для описания изображения человека [2]. Эмпирический поиск дескриптора неочевиден и очень сложен, а сверточные нейронные сети (СНС) позволяют автоматизировать процесс формирования набора признаков. Однако для высокой результативности СНС на этапе обучения необходимо обеспечить выбор наиболее оптимальных гиперпараметров сети, к которым относятся количество итераций обновления весовых коэффициентов, скорость обучения, размер пакета. В алгоритмах реидентификации людей широко используется СНС DenseNet-121 [1]. Следовательно, выбор наиболее оптимальных параметров для данной СНС позволит повысить точность реидентификации с применением ряда алгоритмов на основе DenseNet-121. Архитектура СНС Darknet-53 [3] состоит из Res-блоков, но отличается от ResNet-50 порядком и количеством сверточных слоев на разных уровнях сети, а также

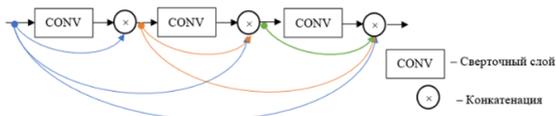
тем, что на выходе каждого нейроны вместо функции активации ReLU, которая традиционно используется в ResNet-50, в DarkNet-53 применяется Leaky-ReLU [4].



**Рис.1. Примеры изображений, отражающие проблему неоднозначности внешнего вида людей при реидентификации:**

- а) разные люди имеют схожий внешний вид;
- б) – изменение визуальных признаков одного и того же человека

Архитектура DenseNet-121 отличается наличием соединений между слоями (рис.2), при которых карты признаков всех предыдущих слоев используются в качестве входных для всех последующих в блоке. Кроме этого, карты признаков не суммируются от слоя к слою, что характерно для ResNet, а конкатенируются. Dense-блоки объединяются в структуры, состоящие из нескольких одинаковых блоков, называемых уровнями сети.



**Рис. 2. Общая структура Dense-блока**

Сложность выбора гиперпараметров обусловлена несколькими факторами. Слишком малый размер пакета не позволяет выявить эффективные признаки для реидентификации и приводит к невозможности высокой обобщающей способности сети. Большой пакет приводит к снижению точности обученной модели, но при этом уменьшает время обучения СНС. Малая скорость обучения приводит к тому, что при фиксированном числе эпох модель не успевает обучиться, следовательно, требуется большее количество эпох, что приводит к увеличению времени тренировки. Увеличивая скорость обучения можно повысить точность, но при этом возникает риск взрывных градиентов, явление, при котором функция потерь начинает очень быстро увеличиваться. Поиск наилучшего сочетания значений скорости обучения, размера пакета и количества эпох зависит от большого числа случайных факторов, решаемой задачи и обучающих данных. Поэтому чаще всего значения гиперпараметров определяются опытным путем.

Для определения наиболее эффективных гиперпараметров СНС DenseNet-121 при реидентификации экспериментальные исследования проводились в два этапа. Для обучения использовался набор данных Market-1501, а для оценки использовались такие критерии как время обучения и метрики точности Rank1, mAP и mINP.

На рис. 3 представлена методика проведения первого этапа экспериментов. На первом этапе необходимо определить наиболее эффективную скорость обучения и размер пакета изображений. Обучение СНС осуществляется с использованием пакетов размером  $b = 2, 4, 8, 16, 32$  и  $64$  изображений. Скорость обучения устанавливается  $\eta = 0,001, 0,002, \dots, 0,009, 0,01, 0,02, \dots, 0,09, 0,01, 0,02, \dots$ . Скорость обучения скрытых слоев меньше, чем скорость обучения классификационного слоя в 10 раз. После 40 эпохи скорость согласно базовой модели алгоритма уменьшается в 10 раз.



**Рис.3. Методология первого этапа экспериментов**

Второй этап предполагает определение количества эпох обучения с учетом изменения размера пакета и скорости в процессе обучения нейронной сети. Начальные значения гиперпараметров были использованы с первого этапа. В процессе обучения размер пакета увеличивался. В табл. 1 представлены экспериментальные результаты, в которых наилучшие показатели точности были получены для трех метрик: Rank1, mAP, mINP.

**Табл.1. Значения метрик оценки точности реидентификации людей с различными гиперпараметрами DenseNet-121**

b	$\eta$	Время обучения	Rank1	mAP	mINP
8	0.04	185m 32s	0,840261	0,642903	0,281465
8	0.05	189m 50s	0,838777	0,648412	0,288388
16	0.06	135m 01s	0,835214	0,639111	0,277537
16	0.08	135m 27s	0,830760	0,626875	0,269123

При двух изображениях в пакете максимальная точность была менее 0,001 для всех метрик, поэтому СНС не обучается. Когда количество изображений в пакете составляет 32 или более и менее 8, это не позволяет повысить точность повторной идентификации по сравнению с данными в таблице 1. При низких

скоростях для достижения сходимости модели требуется большее количество эпох. Использование высокой скорости обучения приводит к более быстрой сходимости, но может привести к взрывным градиентам. Чем больше размер пакета, тем выше может быть скорость обучения. Уменьшение скорости обучения во время обучения СНС снижает значение функции потерь и увеличивает точность повторной идентификации.

На основе анализа полученных экспериментальных данных были исследованы различные подходы к изменению размера пакета и скорости обучения при тренировке нейронной сети. Для первого эксперимента размер пакета удваивался каждые 20 эпох. Скорость обучения изменялась так же, как и для базовой модели, каждые 40 эпох. Для второго и третьего этапов экспериментов пакет изображений увеличивался в 2 раза каждые 30 эпох, но скорость обучения корректировалась аналогично предыдущему этапу. Размер пакета варьировался аналогично предыдущему эксперименту для четвертого этапа. Скорость обучения уменьшалась каждые двадцать эпох в два раза. Для пятого эксперимента размер пакета удваивался после сороковой и шестидесятой эпох, а скорость обучения уменьшалась вдвое каждые десять эпох. Размер пакета увеличивался после пятидесятой и семидесятой эпох для шестого эксперимента. В этом случае мы получили лучшую точность по сравнению с предыдущим экспериментом. Для седьмого эксперимента скорость обучения уменьшалась каждые две эпохи с разными коэффициентами. Наилучшие результаты по точности повторной идентификации были получены для гиперпараметров эксперимента номер восемь:  $b=8, 16; \eta_0=0,04; \eta_0=0,05; \eta_0=0,06; \eta_1=k \cdot \eta_1 - 1, k \in [0,9; 0,99]_{h=0,01}; k=0,98$ . В ходе экспериментов было установлено, что для повышения точности повторной идентификации следует уменьшать скорость обучения после каждой эпохи.

Анализ результатов показывает, что увеличение размера пакета после шестидесятой эпохи позволяет ускорить рост точности повторной идентификации по трем метрикам. Экспериментально определено, что точность повторной идентификации не увеличивается или увеличивается незначительно после восьмидесятой эпохи (рис.4).

В табл. 2 представлены результаты тестирования предложенного подхода на наборах данных DukeMTMC-ReID, MSMT17 и новый набор изображений, рассмотренный в [5] с использованием СНС ResNet-50 и DenseNet-121. Обучение осуществлялось на протяжении 80 эпох.



**Рис.4. Точность повторной идентификации человека на тестовой выборке набора данных Market-1501 для разного количества эпох обучения DenseNet-121**

**Табл.2. Сравнение точности повторной идентификации для базовых и предложенных гиперпараметров**

Набор данных	Гиперпараметры	Rank1	mAP	mINP
DukeMSMT-ReID	Базовые	0,7280	0,5316	0,1588
	Предложенные	0,7500	0,5514	0,1727
PolReID1077	Базовые	0,9033	0,6839	0,2343
	Предложенные	0,9235	0,7060	0,2604
MSMT17	Базовые	0,5498	0,2667	0,0261
	Предложенные	0,5711	0,2788	0,0355

Данные в табл.2 свидетельствуют о том, что экспериментально найденные значения гиперпараметров повышают точность повторной идентификации людей для различных наборов данных с использованием СНС DenseNet-121.

### Литература

1. Чен Х. Повторная идентификация людей в системах видеонаблюдения с использованием глубокого обучения: анализ существующих методов / Х. Чен, С.А. Игнатъева, Р.П. Богущ, С.В. Абламейко // Автоматика и телемеханика. – 2023. – № 5. – 61–112.
2. Игнатъева С.А. Увеличение точности реидентификации людей на основе двухэтапного обучения сверточных нейронных сетей и аугментации / С.А. Игнатъева, Р.П. Богущ // Информатика. – 2023. – № 1. – С. 40 – 54.
3. Redmon, J. YOLOv3: An Incremental Improvement / J. Redmon, A. Farhadi // *ArXiv* abs/1804.02767. – 2018.
4. Maas A. Rectifier non linearities improve neural network acoustic models / A.L. Maas, A.Y. Hannun, A.Y. Ng // *Proceedings of the 30th International Conference on Machine Learning*, Vol. 28.
5. Игнатъева С.А. Формирование обучающей выборки для сверточных нейронных сетей при реидентификации людей по видеоданным / С.А. Игнатъева, Р.П. Богущ // Доклады БГУИР. – 2023. – № 3. – С. 87 – 95.