

Развитие технологий распознавания на основе дополнительных признаков

А. А. Зацаринный¹, А. А. Карандеев², А. Е. Маслов¹, В. П. Осипов², Н. Ю. Апальков³

¹ Федеральный исследовательский центр «Информатика и управление» РАН, Москва, Россия

² Федеральный исследовательский центр «Институт прикладной математики им. М.В. Келдыша» РАН, Москва, Россия

³ Российский экономический университет имени Г. В. Плеханова, Москва, Россия

Аннотация. В статье рассмотрены различные технологии распознавания изображений и предложены способы их улучшения путем исследования дополнительных признаков. В частности, представлен новый подход, который способствует улучшению решения задачи распознавания путем использования точек Харриса в качестве дополнительных признаков на изображениях. Это существенно повышает точность классифицирующей модели распознавания. Значимость данного подхода заключается в его способности улучшить возможности системы распознавания в обнаружении и выделении ключевых особенностей объектов, что в конечном счете приводит к более надежным и эффективным результатам при анализе, обработке и классификации данных, а также повышает устойчивость модели. Благодаря этим улучшениям, данная технология распознавания может быть успешно применена в различных областях, где требуется высокая точность и надежность при распознавании информации, например в сфере медицины, классификации транспортных средств и т.д.

Ключевые слова: распознавание образов, нейронные сети, детектор Харриса, устойчивость модели, искусственный интеллект.

DOI 10.14357/20718632240107

EDN CZSVDM

Введение

Актуальность данного исследования обусловлена постоянно растущей потребностью в эффективных и надежных системах распознавания изображений в различных сферах жизни, таких как медицина, безопасность, автоматизация производственных процессов, управление транспортом и многих других [1-3].

В последние десятилетия наблюдается стремительное развитие технологий распознавания, играющих важную роль во многих сферах нашей жизни. Распознавание объектов, текста, лиц, жестов и речи стало неотъемлемой частью современного информационного общества

и создает условия для эффективной автоматизации многих процессов управления и принятия решений [4]. Однако вопреки значительному прогрессу в области распознавания, существующие методы все еще сталкиваются с *вызовами*, связанными с устойчивостью к разнообразным помехам и значительным изменениям во внешней среде [5-7].

Множество новых исследований направлены на изучение информативности различных типов признаков объекта и внешней среды в процессе распознавания. Это могут быть как контекстные признаки, связанные с окружением объекта или события, так и дополнительные данные, такие как тепловое изображение при распознавании

лиц, а также динамические признаки, отображающие изменение объектов во времени [8, 9]. Использование таких дополнительных признаков позволяет повысить устойчивость алгоритмов распознавания к разнообразным помехам в исходных данных, включая изменения в освещении, вариации масштаба, наличие стороннего шума и других возмущений.

Для распознавания видео и изображений часто используется семантическая сегментация и выделение ключевых точек [10]. Это позволяет распознавать как наличие объекта на изображении или видео, так и его местоположение с помощью сверточных нейронных сетей.

Благодаря развитию машинного обучения и глубоких нейронных сетей, эти методы становятся все более эффективными и способными обрабатывать большие объемы данных с высокой точностью.

В данной статье основное внимание уделено исследованию возможности использования дополнительной информации, неявно закодированной в изображениях, для повышения устойчивости искусственных нейронных сетей. Предлагаемый подход к развитию технологий распознавания предполагает поиск и использование дополнительных признаков, которые уже встроены в сами изображения или данные, подвергнутые предварительной обработке. Это позволяет извлекать ценную информацию из контекста или свойств самого объекта, что существенно расширяет способности модели распознавания.

1. Методические предпосылки повышения устойчивости распознавания объектов

В монографии [11] для повышения устойчивости НС предлагается использовать распознавание контуров изображения. При этом учитывается то обстоятельство, что при классификации изображения НС делают акцент на текстуре изображения, практически не затрагивая контур. Человек же, напротив, принимает решения, руководствуясь главным образом формой объекта.

В работе [12] исследователи создали новый набор обучающих данных со «стилизированными»

изображениями, метки которых соответствовали форме объектов. Это позволило повысить устойчивость классификаторов к вносимым в изображения искажениям, поскольку большинство искажений изменяет лишь текстуру изображения, практически не затрагивая контур объекта. Такого рода исследования показывают необходимость привлечения дополнительных информативных признаков для повышения вероятности правильного распознавания объектов.

Оценку повышения вероятности корректной работы НС можно получить на основе формулы Байеса:

$$P(A|B) = \frac{P(B|A)P(A)}{P(B)},$$

где: $P(A|B)$ — условная вероятность того, что событие "А" произошло, при условии, что произошло событие "В",

$P(B|A)$ — условная вероятность того, что событие "В" произошло, при условии, что произошло событие "А",

$P(A)$ — начальная вероятность события "А",

$P(B)$ — вероятность события "В".

Формула дает возможность дать новую оценку вероятности события "А" при условии, что произошло событие "В". То есть, формула Байеса поможет уточнить наши представления о вероятности адекватного распознавания объекта в случае распознавания дополнительных информативных признаков. На этом основано предложение по поиску дополнительных информативных признаков распознаваемых объектов, которые можно использовать для дополнительного обучения НС. В качестве таких информативных признаков могут быть использованы локальные дескрипторы изображений [6]. Этот подход может помочь сделать обучение НС более устойчивыми и улучшить обобщающую способность модели.

Так, можно привести пример использования формулы Байеса, с помощью которой можно определить вероятность распознавания объекта в зависимости от дополнительных данных или признаков (Рис. 1). На данном рисунке показатель Alpha отображает интенсивность наложения помех на изображение, а верхняя линия показывает точность модели распознавания объекта без введения дополнительных данных (без учета показателя Alpha). Предположим, что

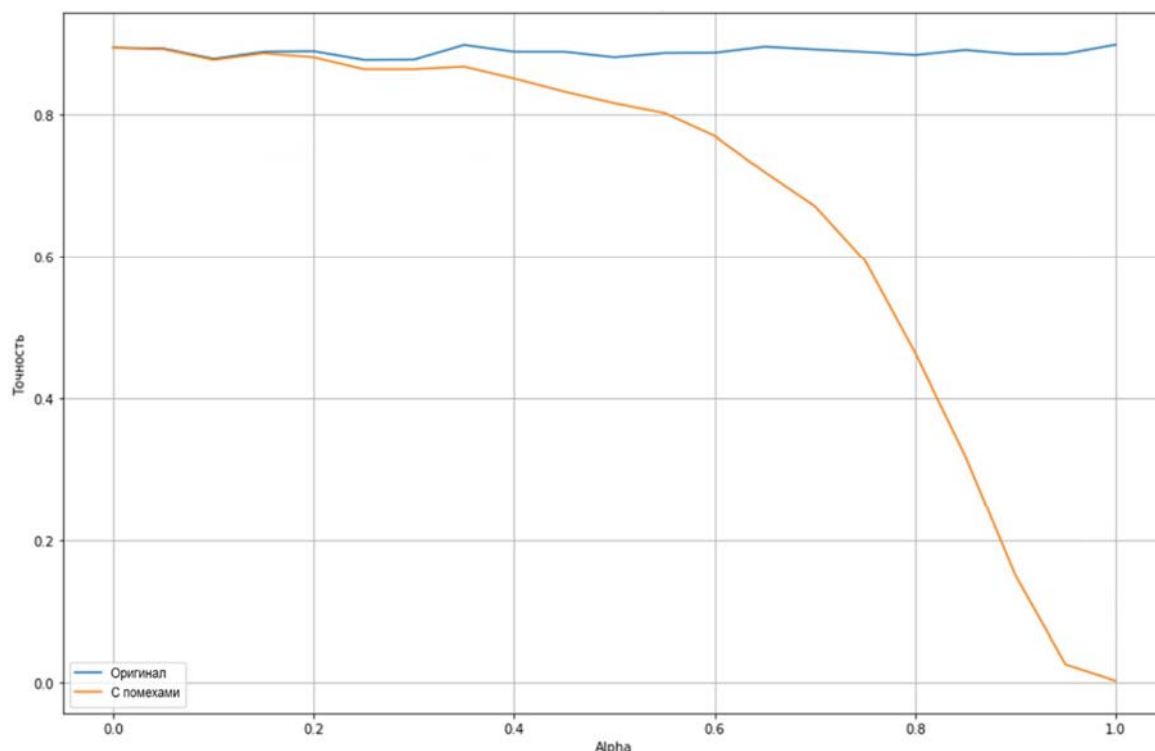


Рис. 1. Вероятность распознавания объекта в зависимости от интенсивности вредоносного воздействия

исследования влияния некоторых видов воздействия на вероятности адекватного распознавания различных предварительно обученных НС показали, что наблюдаемая вероятность правильного распознавания объекта на изображении, после вредоносного воздействия снижается до практически неприемлемых значений $P(A) = 0.6$ и ниже. Предположим, что мы ввели в рассмотрение дополнительный информативный признак B . Вероятность распознавания этого признака при правильном распознавании объекта A будет приближаться к единице, предположим для иллюстрации $P(B|A) = 0.9$. Полная вероятность распознавания дополнительного признака B будет равна сумме вероятностей его правильного распознавания как при наличии события A (адекватного распознавания объекта), так и его отсутствии (объект A не распознан) $P(B) = 0.9 \cdot 0.6 + 0.1 \cdot 0.4 = 0.58$.

После использования формулы Байеса получим вероятность распознавания объекта в зависимости от заданных параметров: $P(A|B) = 0.93$.

Таким образом, использование дополнительных информативных признаков поможет

существенно повысить вероятностные предсказания. Это может помочь более осознанно использовать результаты работы НС и принимать решения, основанные на более высокой степени уверенности в применяемой модели, основанной на оптимальных наборах информативных признаков и исходных данных для обучения моделей.

В качестве объекта исследования были выбраны изображения самолетов из набора данных APL_Extended¹, а дополнительными признаками выступают характерные для самолетов особые точки - углы, обнаружение которых было осуществлено методом на основе алгоритма обнаружения углов Харриса, который заключается в поиске особых (угловых) точек, в окрестности которых имеются границы в нескольких направлениях [13]. Применение метода должно повысить точность модели за счет введения дополнительных признаков. Помимо этого, метод будет применен к изображениям, которые были подвергнуты операции размытия (blur), для демонстрации возможности повышения устойчивости модели. Так, в рамках алгоритма Харриса,

¹ <https://meta-album.github.io/datasets/APL.html>

для каждой точки изображения определяется положительно-полуопределенная симметричная матрица $M_I = M_I(x)$, где x - точка внутри изображения:

$$M_I = \nabla I \nabla I^T = \begin{bmatrix} I_x \\ I_y \end{bmatrix} \begin{bmatrix} I_x I_y \end{bmatrix} = \begin{bmatrix} I_x^2 & I_x I_y \\ I_x I_y & I_y^2 \end{bmatrix}, \quad (1)$$

где: ∇I - вектор градиента изображения, содержащий производные I_x и I_y .

Матрица M_I имеет ранг 1, собственные значения равны $\lambda_1 = |\nabla I|^2$ и $\lambda_2 = 0$. Обозначим W матрицу весов. Операция поэлементной свертки

$$\overline{M_I} = W * M_I \quad (2)$$

дает локальное усреднение $\overline{M_I}$ по соседним пикселям.

Для выявления особых точек без вычисления собственных значений в работе [12] ввели индикаторную функцию:

$$R = \det(\overline{M_I}) - \kappa \text{trace}(\overline{M_I})^2, \quad (3)$$

где: $\det(\overline{M_I}) = \lambda_1 \lambda_2$,

$\text{trace}(\overline{M_I}) = \lambda_1 + \lambda_2$,

κ – эмпирически определяемая константа в отрезке $[0.04, 0.06]$.

Таким образом, величины этих собственных значений определяют, является ли граница области гладкой, или же содержит угол, ребро:

- когда $|R|$ мало, что бывает при малых λ_1 и λ_2 , граница области гладкая (угол отсутствует);

- при $R < 0$, что бывает при $\lambda_1 \gg \lambda_2$ или наоборот, область содержит ребро;

- когда R велико, что происходит, когда λ_1 и λ_2 велики и $\lambda_1 \sim \lambda_2$, область содержит угол.

Для избавления от весовой постоянной κ в качестве индикатора используется частное

$$\frac{\det(\overline{M_I})}{\text{trace}(\overline{M_I})^2} \quad (4)$$

Таким образом, в ходе применения алгоритма Харриса, на основе исходного изображения, формируется дополнительное изображение, которое содержит в себе ключевые углы на изображении. В эти участки в дальнейшем может кодироваться та или иная информация в зависимости от поставленной задачи и необходимых результатов работы модели.

2. Результаты вычислительных экспериментов

В качестве входного набора данных для обучения ИНС, из набора данных APL Extended было использовано 1960 изображений самолетов, состоящие из 5 классов (Рис. 2):

После вычисления особых точек для каждого изображения из набора данных, формируется маска, которая в дальнейшем наносится на изображение с определенным уровнем прозрачности, который практически невидим для человеческого глаза. Для каждого класса изображений

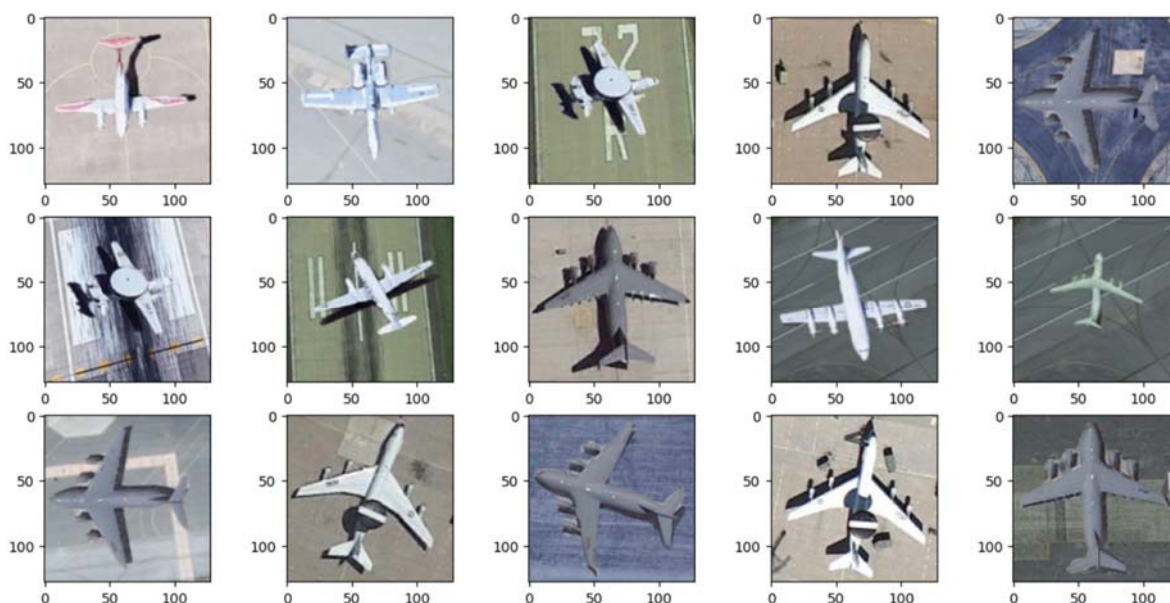


Рис. 2. Пример изображений из исследуемого набора данных

используется уникальный цвет, который несет в себе дополнительную информацию относительно принадлежности к тому или иному классу.

На Рис. 3 представлены маски, состоящие из особых точек, а также представлены соответствующие изображения с наложенными масками (прозрачность уменьшена для визуализации).

Результаты эксперимента показали, что применение метода к набору данных увеличивает точность и повышает устойчивость модели распознавания. Так, точность модели, обученной на исходных изображениях составляет 87,5%. Точность модели, обученной на исходных изображениях после применения метода к исследуемому набору данных составляет уже 92,7%.

В качестве эксперимента, было принято решение наложить те же самые маски (Рис. 3) на размытые изображения (Рис. 4).

Таким образом, была смоделирована ситуация, когда часть данных могла повредиться при передаче, или качество изображения было ухудшено вследствие его сжатия. В данном эксперименте точность модели, обученной на размытых изображениях, составляет 70,4% (Рис. 5), ядро размытия равное трём). Точность модели, обученной на исходных изображениях, после

применения метода к исследуемому набору данных составляет уже 82,6%.

Заключение

1. Предложен и апробирован способ повышения устойчивости модели распознавания изображения, основанный на введении дополнительной информации во входные данные. Эффективность способа продемонстрирована на задаче распознавания изображений, где в качестве дополнительных признаков использовались точки Харриса, построенные поверх изображений и размеченные для обучения вручную;

2. По результатам экспериментальных исследований выявлено, что предложенный способ позволяет:

- существенно повысить точность распознавания данных с учетом их восстановления и фрагментации;
- обеспечить условия повышения устойчивости искусственных нейронных сетей к вредоносному или случайному воздействию на входные данные на основе введения в установленные ключевые точки закодированной дополнительной информации, это позволит лучше классифицировать объекты на изображениях;

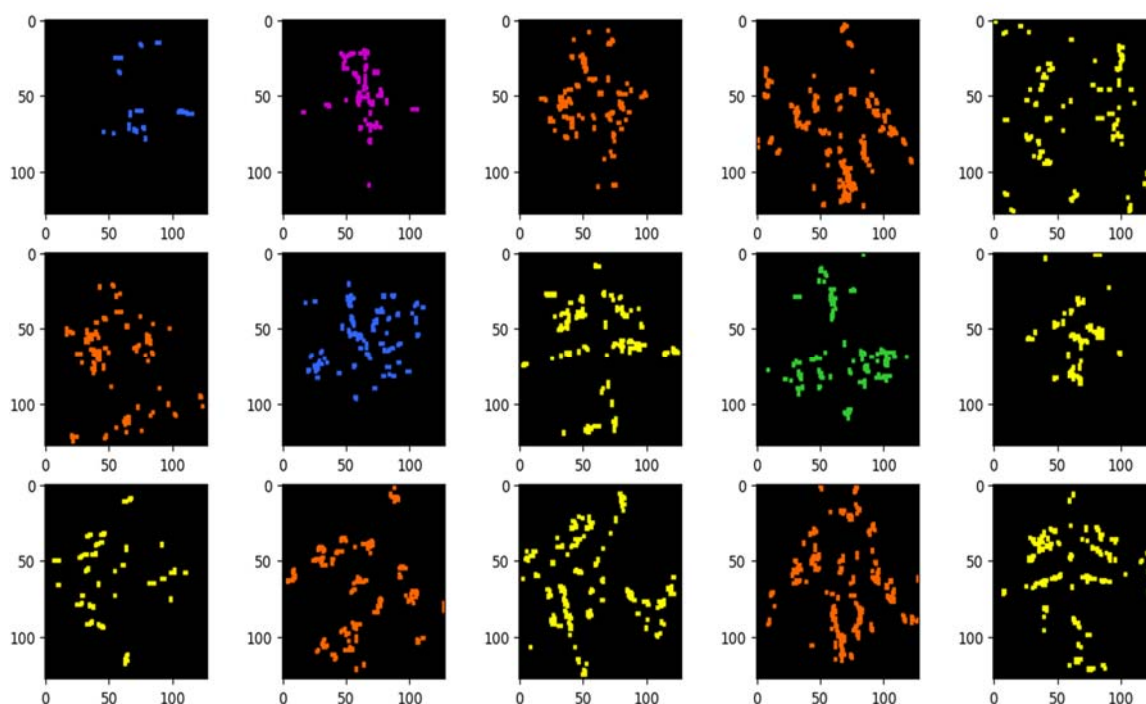


Рис. 3. Пример масок, состоящих из особых точек

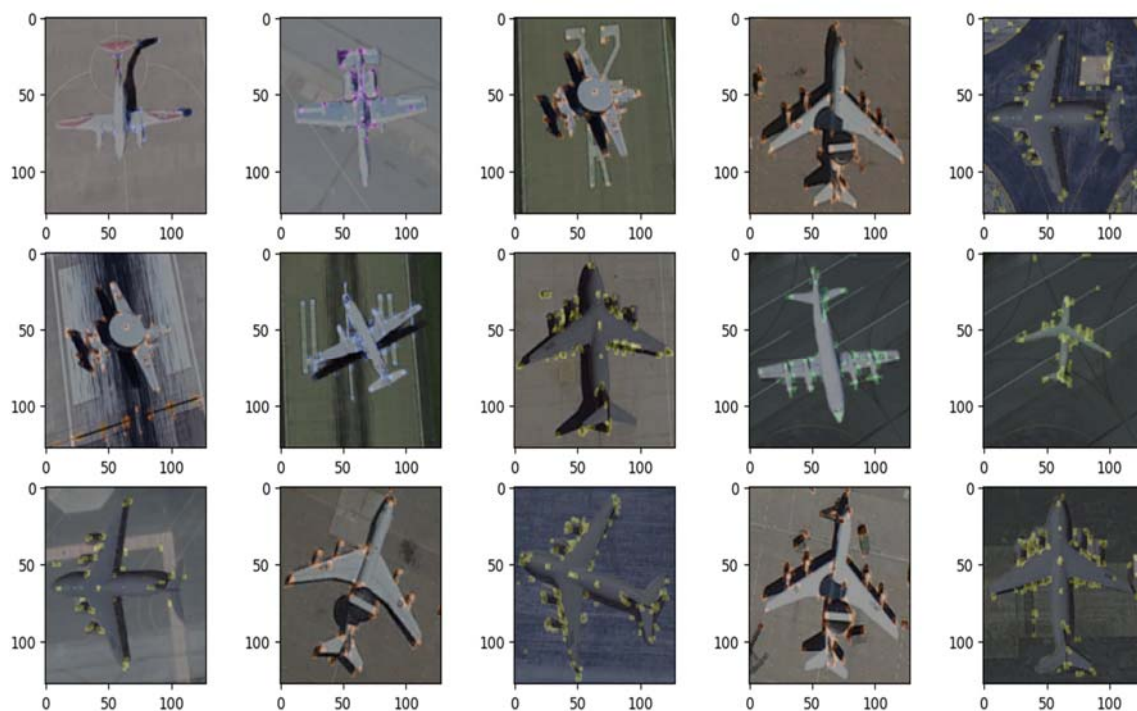


Рис. 4. Пример изображений после наложения особых точек

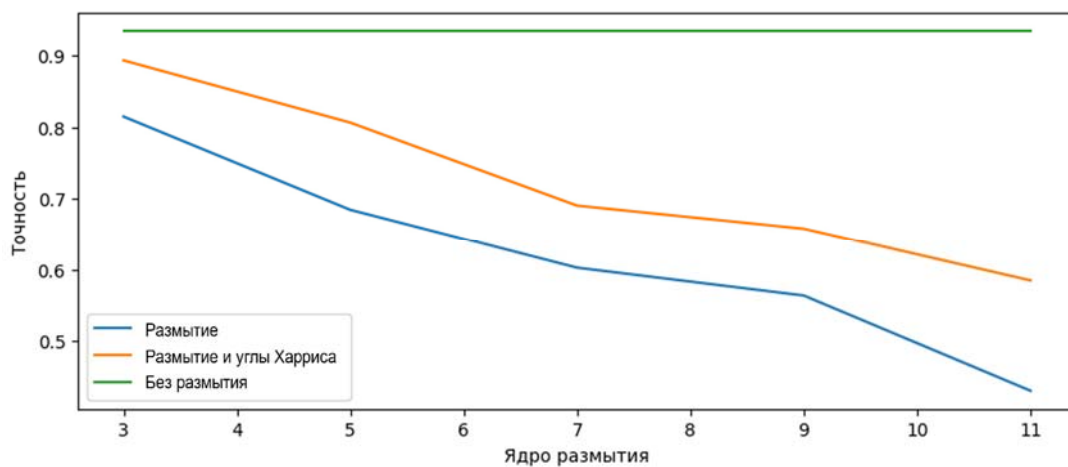


Рис. 5. Точность модели с использованием размытия при добавлении дополнительных признаков

3. Показано, что использование и развитие предлагаемого способа является перспективным направлением изыскания эффективных методов повышения устойчивости моделей и повышения точности распознавания изображений в условиях интенсивных негативных внешних воздействий;

4. Предполагаются дальнейшие исследования по развитию предлагаемого способа применительно к различным задачам распознавания.

Литература

1. Yue X., Lyu B., Li H., Meng L., Furumoto K. Real-time medicine packet recognition system in dispensing medicines for the elderly // Measurement: Sensors. 2021.
2. Mahami H., Ghassemi N., Darbandy M. T., Shoeibi A., Hussain S., Nasirzadeh F., Alizadehsani R., Nahavandi D., Khosravi A., Nahavandi S. Material Recognition for Automated Progress Monitoring using Deep Learning Methods // Computer Vision and Pattern Recognition. 2021.
3. Rinne M., Bagheri M., Tolvanen T., Hollmén J. Automatic Recognition of Public Transport Trips from Mobile Device

- Sensor Data and Transport Infrastructure Information // International Workshop on Personal Analytics and Privacy. 2017.
4. Shadiev R., Zhang Z. H., Wu T.-T., Huang Y.M. Review of Studies on Recognition Technologies and Their Applications Used to Assist Learning and Instruction // Educational Technology & Society. 2020. № 4. pp 59-74.
 5. Waelen R.A. The struggle for recognition in the age of facial recognition technology // AI and Ethics. 2023. № 3. pp 215–222.
 6. Солям Я. Э., Программирование компьютерного зрения на языке Python/ - М.ДМК Пресс, 2016. – 312 с.
 7. Селянкин В.В. Компьютерное зрение. Анализ и обработка изображений – Санкт Петербург: Лань, 2021.- 152 с.
 8. Papastratis I., Dimitropoulos K., Daras P. Continuous Sign Language Recognition through a Context-Aware Generative Adversarial Network // Sensors. 2021. № 7. 2437.
 9. Sun Z., Ke Q., Rahmani H., Bennamoun M., Wang G., Liu J. Human Action Recognition from Various Data Modalities: A Review // IEEE transactions on pattern analysis and machine intelligence. 2022.
 10. Lui X., Deng Z., Yang Y. Recent progress in semantic image segmentation // Artificial Intelligence Review. 2018. № 2. pp 1089-1106.
 11. Уорр Кэти, Надежность нейронных сетей: укрепляем устойчивость ИИ к обману. — СПб.: Питер, 2021. — 272 с.
 12. Geirhos R. et al. ImageNet-Trained CNNs Are Biased Towards Texture; Increasing Shape Bias Improves Accuracy and Robustness, 2019.
 13. Harris C. G., Stephens M. J. A Combined Corner and Edge Detector // Alvey Vision Conference. 1988.

Зацаринный Александр Алексеевич. ФИЦ ИУ РАН. Главный научный сотрудник, доктор технических наук, профессор. Область научных интересов: интегрированные информационно-телекоммуникационные системы и сети, искусственный интеллект, процессы информатизации управления. E-mail: AZatsarinny@ipiran.ru

Карандеев Александр Андреевич. Федеральный исследовательский центр «Институт прикладной математики им. М.В. Келдыша» РАН. Научный сотрудник, кандидат технических наук. Область научных интересов: математическое моделирование, искусственный интеллект, нейронные сети, принятие решений в условиях неопределенности. E-mail: KarAlex755@gmail.com

Маслов Алексей Евгеньевич. ФИЦ ИУ РАН. Научный сотрудник. Область научных интересов: компьютерное зрение (CV), анализ и обработка естественного языка (NLP), анализ больших массивов данных, разработка распределенных программных систем. E-mail: amaslov@frcsc.ru

Осипов Владимир Петрович Федеральный исследовательский центр «Институт прикладной математики им. М.В. Келдыша» РАН. Ведущий научный сотрудник, кандидат технических наук. Область научных интересов: моделирование сложных систем, искусственный интеллект, принятие решений в условиях неопределенности. E-mail: osipov@keldysh.ru

Апальков Никита Юрьевич Российский экономический университет имени Г. В. Плеханова. Студент. Область научных интересов: компьютерный анализ естественного языка, искусственный интеллект, методы интеллектуального информационного поиска, анализ больших массивов данных. E-mail: nikita_apalkov@mail.ru

Development of Technologies Based on Additional Properties

A. A. Zatsarinny¹, A. A. Karandeev¹¹, A. E. Maslov¹, V. P. Osipov¹¹, N. Y. Apalkov¹¹¹

¹ Federal Research Center "Computer Science and Control of Russian Academy of Sciences", Moscow, Russia

¹¹ Federal Research Center "Keldysh Institute of Applied Mathematics of Russian Academy of Sciences", Moscow, Russia

¹¹¹ Plekhanov Russian University of Economics, Moscow, Russia

Abstract. The article discusses various image recognition technologies and proposes methods to enhance them by exploring additional features. In particular, a new approach is introduced that contributes to improving image recognition by using Harris corners as additional features in images. This significantly enhances the accuracy of the recognition classification model. The significance of this approach lies in its ability to enhance the recognition system's capabilities in detecting and highlighting key object features, ultimately leading to more reliable and efficient results in data analysis, processing, and classification. It also increases the model's robustness. Thanks to these improvements, this image recognition technology can be successfully applied in various fields where high accuracy and reliability are required in information recognition, such as medicine, vehicle classification, and more.

Keywords: pattern recognition, neural networks, Harris detector, model stability, artificial intelligence.

DOI 10.14357/20718632240107

EDN CZSVDM

References

1. Yue X., Lyu B., Li H., Meng L., Furumoto K. Real-time medicine packet recognition system in dispensing medicines for the elderly // *Measurement: Sensors*. 2021.
2. Mahami H., Ghassemi N., Darbandy M. T., Shoeibi A., Hussain S., Nasirzadeh F., Alizadehsani R., Nahavandi D., Khosravi A., Nahavandi S. Material Recognition for Automated Progress Monitoring using Deep Learning Methods // *Computer Vision and Pattern Recognition*. 2021.
3. Rinne M., Bagheri M., Tolvanen T., Hollmén J. Automatic Recognition of Public Transport Trips from Mobile Device Sensor Data and Transport Infrastructure Information // *International Workshop on Personal Analytics and Privacy*. 2017.
4. Shadiev R., Zhang Z. H., Wu T.-T., Huang Y.M. Review of Studies on Recognition Technologies and Their Applications Used to Assist Learning and Instruction // *Educational Technology & Society*. 2020. № 4. pp 59-74.
5. Waelen R.A. The struggle for recognition in the age of facial recognition technology // *AI and Ethics*. 2023. № 3. pp 215–222.
6. Solem Ya. E., *Programming computer vision in Python/ - M.DMK Press*, 2016. – 312 p
7. Selyankin V.V. *Computer vision. Image analysis and processing – Saint Petersburg: Lan*, 2021.- 152 p.
8. Papastratis I., Dimitropoulos K., Daras P. Continuous Sign Language Recognition through a Context-Aware Generative Adversarial Network // *Sensors*. 2021. № 7. 2437.
9. Sun Z., Ke Q., Rahmani H., Bennamoun M., Wang G., Liu J. Human Action Recognition from Various Data Modalities: A Review // *IEEE transactions on pattern analysis and machine intelligence*. 2022.
10. Lui X., Deng Z., Yang Y. Recent progress in semantic image segmentation // *Artificial Intelligence Review*. 2018. № 2. pp 1089-1106.
11. Warr Katie, *Reliability of neural networks: strengthening the resistance of AI to deception. — St. Petersburg: Peter*, 2021. — 272 p.
12. Geirhos R. et al. ImageNet-Trained CNNs Are Biased Towards Texture; Increasing Shape Bias Improves Accuracy and Robustness, 2019.
13. Harris C. G., Stephens M. J. A Combined Corner and Edge Detector // *Alvey Vision Conference*. 1988.

Zatsarinny Alexander A. Doctor of Science in technology, professor, principal scientist, Federal Research Center “Computer Science and Control of Russian Academy of Sciences”, Moscow, Areas of scientific interest: integrated information and telecommunications systems and networks, artificial intelligence, processes of management informatization. E-mail: AZatsarinny@ipiran.ru

Karandeev Alexander A. PhD in Engineering sciences, Federal Research Center "Keldysh Institute of Applied Mathematics of Russian Academy of Sciences", Moscow, Researcher. Areas of scientific interest: mathematical modeling, artificial intelligence, neural networks, decision-making under uncertainty. E-mail: KarAlex755@gmail.com

Maslov Alexey E. Federal Research Center “Computer Science and Control of Russian Academy of Sciences”, Moscow, Researcher. Research interests: computer vision (CV), natural language analysis and processing (NLP), analysis of large data sets, development of distributed software systems. E-mail: amaslov@frccsc.ru.

Osipov Vladimir P. PhD in Engineering sciences, Federal Research Center "Keldysh Institute of Applied Mathematics of Russian Academy of Sciences" of the Russian Academy of Sciences, Moscow, leading researcher. Areas of scientific interest: modeling of complex systems, artificial intelligence, decision-making under uncertainty. E-mail: osipov@keldysh.ru

Apalkov Nikita Y. Plekhanov Russian University of Economics, Moscow, Master's degree student in Applied Informatics. Areas of scientific interest: computer analysis of natural language, artificial intelligence, methods of intelligent information retrieval, analysis of large data arrays. E-mail: nikita_apalkov@mail.ru