

# Сравнительный анализ методов машинного обучения для классификации объектов на последовательных изображениях звёздного неба

А. А. Щукин, М. Д. Поляк

Санкт-Петербургский государственный университет аэрокосмического приборостроения

Antonovka2310@yandex.ru, markpolyak@gmail.com

**Аннотация.** В данной работе решается задача классификации объектов на видеоряде звёздного неба, заключающаяся в разделении пятен на фоновые звёзды и движущиеся объекты. Основное внимание уделяется сравнительному анализу двух подходов: метода опорных векторов (SVM) и нейросетевого метода. В качестве признаков для SVM использовалась траектория движения объектов, нейросеть обрабатывает траектории объектов и выделяет признаки самостоятельно.

**Ключевые слова:** метод опорных векторов; U-Net; классификация

## I. ВВЕДЕНИЕ

В последние десятилетия наблюдается повышенная заинтересованность к изучению и развитию космической отрасли, в том числе развитие оптических средств наблюдения. Одной из актуальных задач в данной области является автоматическая классификация объектов на видеоряде звёздного неба, включающая выделение и идентификацию как фоновых звёзд, так и движущихся объектов различной природы — спутников, метеоров и фрагментов космического мусора.

Ключевая сложность данной задачи заключается в точном определении объектов на зашумлённом и динамически сложном фоне звёздного неба. На первичном этапе анализа видеоряда пятна — пока что не классифицированные объекты или звёзды — могут иметь сложную траекторию движения, различные скорость и паттерны движения. Для решения этой задачи необходимо выделить признаки, наилучшим образом отличающие звёзды от космических объектов, разработать устойчивый к изменениям алгоритм, решающий задачу классификации.

В данной работе рассматривается задача классификации объектов на видеоряде звёздного неба, включающая выделение и идентификацию как фоновых звёзд, так и движущихся объектов различной природы — спутников, метеоров и фрагментов космического мусора.

Целью исследования является создание решения для автоматической классификации таких объектов. В работе осуществляется сравнительный анализ двух подходов: метода опорных векторов (SVM) и нейросетевого подхода на базе архитектуры U-Net. Исследование направлено на оценку эффективности этих методов в условиях синтетически сгенерированных видеоданных, моделирующих реальные снимки звездного неба, с целью выявления их сильных и слабых сторон, а также определения перспективности дальнейшего использования в системах мониторинга космического пространства.

## II. ОПИСАНИЕ МЕТОДОВ МАШИННОГО ОБУЧЕНИЯ

### A. Метод опорных векторов (SVM)

SVM (Support Vector Machines) — алгоритм машинного обучения, который используется для решения задач классификации. За последние несколько лет SVM стал одним из лучших методов классификации, во многих случаях превосходящий по точности другие модели [1–4].

Основная задача метода заключается в нахождении такой гиперплоскости, которая в большей степени сохраняет исходное распределение точек в признаковом пространстве по возможности без дополнительных предположений.

Ядра модели — математические функции, которые определяют поведение гиперплоскости относительно разных типов входных данных. Представлены наиболее популярные ядра для классификации по двум векторам признаков  $X_i$  и  $X_j$ :

- Линейное ядро:  $K = x_i * x_j$  — простейший вид ядра, является скалярным произведением признаков, используется, когда классы являются линейно разделимыми;
- Полиномиальное ядро:  $K = (\gamma \cdot x_i * x_j + r)^d$ , где  $r$  — свободный член,  $d$  — степень полинома,  $\gamma$  — гиперпараметр; такое ядро учитывает более сложные взаимодействия признаков, например квадратичные или кубические зависимости;
- Сигмовидное ядро:  $K = \tanh(\gamma \cdot x_i * x_j + r)$ ;
- Радиально-базисное ядро (RBF-ядро):  $K = \exp(-\gamma \|x_i - x_j\|^2)$  — наиболее популярное ядро, которое работает с линейно-неразделимыми данными.

Несмотря на гибкую настройку и все преимущества данного подхода, SVM-классификатор требует обязательной обработки входных признаков. Данный алгоритм чувствителен к масштабированию и выбросам, поскольку работает на основе расстояний и скалярных произведений между объектами. Поэтому перед обучением модели требуется стандартизация или нормализация данных. Кроме того, наличие аномальных значений или шумовых выбросов может существенно повлиять на положение опорных векторов, что снижает обобщающую способность модели.

### В. Нейросетевой подход

Данный подход требует создания и обучения нейросетевой модели для решения задачи классификации. Преимущества данного подхода заключается в том, что нейросеть сама выделяет признаки для классификации [5].

В данной работе использовалась свёрточная нейронная сеть с архитектурой U-Net [6, 7]. Она названа так из-за своей симметричной U-образной формы, отражающей структуру сети: энкодер (сжимающая часть), декодер (восстанавливающая часть) и ядро принятия решений.

На рис. 1 представлена общая схема модели U-Net, которая имеет три этапа свёртки. На вход модели подаётся одноканальное изображение (изображение в градациях серого, где значение яркости каждого пикселя лежит в диапазоне от 0 до 255) размером  $N \times M$  пикселей.

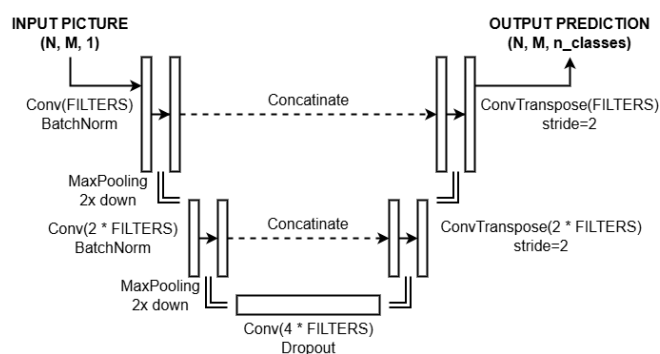


Рис. 1. Схематичная визуализация архитектуры U-Net

Энкодер состоит из нескольких блоков, каждый блок состоит из одного свёрточного слоя и одного слоя MaxPooling, который сжимает изображение в 2 раза. Энкодер необходим, чтобы модель смогла распознать, что изображено на входной картинке. На каждом блоке через фильтры определяются разные по значимости признаки.

На рис. 2 изображён пример свёртки изображения. Из исходного изображения выделяется матрица заданной размерности и умножается на матрицу-фильтр, определитель результирующей матрицы будет пикселем нового изображения. Используя соответствующие фильтры можно выделять нужные паттерны изображения [8].

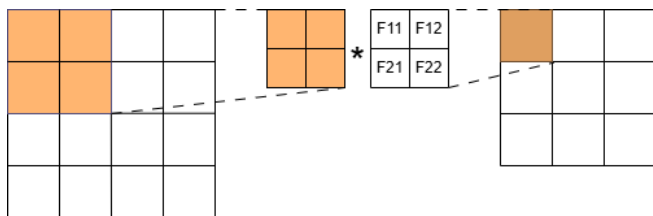


Рис. 2. Пример свёртки изображения

Ядро является центром принятия решения в данной модели. На вход данного слоя поступает наиболее сжатое изображение, но при этом его обрабатывает наибольшее количество фильтров. Для регуляции данного слоя необходимо подключить Dropout, чтобы избежать переобучения модели.

Декодер необходим для развёртки сжатого изображения и создания классификатора. Чтобы модель смогла корректно отобразить каждый объект, необходимо передавать изображение из “параллельного” блока из Энкодера.

На выходе модели формируется матрица той же размерности  $N \times M$ , однако значениями элементов этой матрицы являются вероятности принадлежности соответствующего исходного пикселя изображения заданному классу.

### III. ПРАКТИЧЕСКИЙ РАЗДЕЛ

#### А. Описание датасета

В данной работе использовались искусственно-смоделированные видеоряды звёздного неба, на которых присутствуют звёзды и космические объекты. Видео отличаются между собой скоростью движения фоновых звёзд. К каждому видеоряду прилагается файл в формате MATLAB, в котором для каждого кадра обозначены координаты центроидов звёзд и объектов [9]. Датасет создан с помощью динамического симулятора звёздного поля Центра космических полётов NASA Marshall. В отличие от реальных снимков [10], сгенерированный видеоряд чище и менее зашумлён. На рисунке 3 представлен реальный снимок звёздного неба, сделанный Sapienza Space debris Observatory Network.

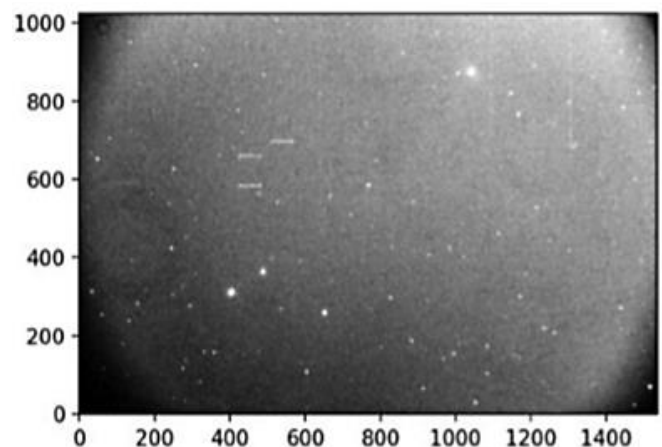


Рис. 3. Снимок звёздного неба, предоставленный Sapienza Space debris Observatory Network

#### В. Реализация SVM-классификатора

Как было сказано выше, SVM-классификатор работает со статичными данными, поэтому первоначально необходимо вычленить признаки, по которым модель будет классифицировать пятна. В качестве признаков были выделены различия в скорости и угле движения каждого пятна относительно других пятен. Основная идея заключается в том, что космические объекты, в отличие от фоновых звёзд, расположены ближе к наблюдателю, поэтому они имеют другую траекторию движения. Для обучения классификатора использовались примеры из разных видеорядов, чтобы алгоритм был более устойчив к разным фоновым движениям звёзд.

Для извлечения признаков непосредственно из видеоряда использовался алгоритм субструктуры фона MOG2, который на основе яркости и перемещения каждого конкретного пикселя определяет пятна и

отличает их от фона [11, 12]. Далее вычисляется оптический поток изображения методом Лукаса-Канаде, который считает вектор перемещения каждого пятна из кадра  $t$  в позицию на кадре  $t+1$  [13]. После классификации каждый объект помечается нужным образом для визуализации результатов. На рис. 4 представлена блок-схема алгоритма подготовки изображения к классификации, используя алгоритм SVM.

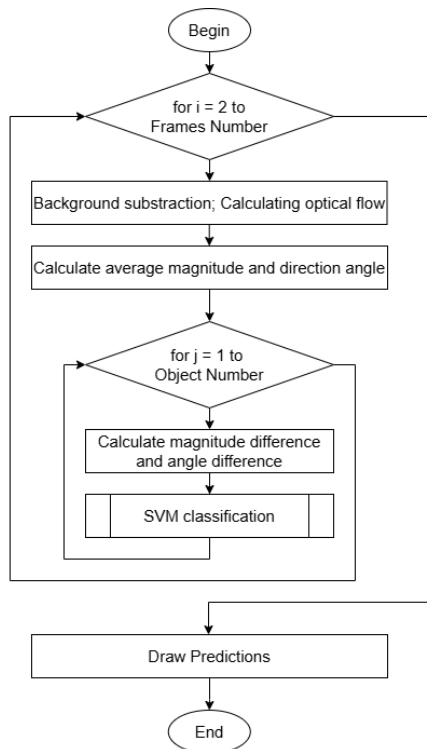


Рис. 4. Схема обработки видеоряда для SVM-классификатора

На рис. 5 представлен пример классификации пятен. Фиолетовым цветом помечаются неизвестные пятна, которые пока что нельзя классифицировать ввиду недостатка данных, серым отмечены звёзды, а зелёным цветом в рамке отмечены объекты.



Рис. 5. Пример классификации пятен

В данной работе использовался SVM-классификатор с RBF-ядром, который показал следующие метрики на тестовой выборке.

ТАБЛИЦА I. МЕТРИКИ SVM-КЛАССИФИКАТОРА

| Название классов | precision | recall | f1-score |
|------------------|-----------|--------|----------|
| 0 – Звёзды       | 0.98      | 0.99   | 0.98     |
| 1 – Объекты      | 0.99      | 0.98   | 0.98     |

### С. Реализация нейронной сети U-Net

Идея использовать нейронную сеть объясняется тем, что, в отличие от SVM-классификатора, нейросеть сама определяет ключевые признаки, по которым будет различать объекты. Доверяя нейронной сети самостоятельное выделение признаков, можно исключить влияние человеческого фактора и выявить более значимые признаки.

U-Net не является рекуррентной моделью, поэтому в качестве входных данных подаётся Motion History Image (MHI). Данное изображение будет хранить в себе как нынешнее, так и прошлые позиции пятен, формируя их траектории движения [14]. На рис. 6 представлен фрагмент MHI.



Рис. 6. Пример MHI

На рис. 7 представлена блок-схема подготовки изображения видеоряда к классификации, используя U-Net.

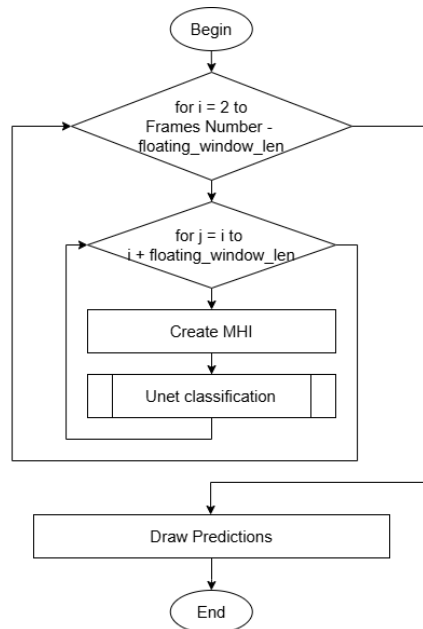


Рис. 7. Схема обработки видеоряда для U-Net

Так как выбранная нейросетевая модель классифицирует каждый пиксель изображения, наблюдается сильный дисбаланс классов. Для решения данной проблемы к каждому кадру дорисовывались копии маски объектов, отражённые по горизонтали, по вертикали и по обеим осям. Также для классов были

обозначены веса, которые увеличивают значимость предсказаний звёзд и объектов.

В работе использовалась трёхуровневая модель U-net с 32 фильтрами на первом свёрточном блоке, 64 и 128 фильтрами на последующих блоках. Всего модель имеет 203939 настраиваемых параметров, обучалась на выборке из одного видео с 696 примеров в обучающей и 174 в тестовой выборке в течении 5 эпох с размером батчей 8 примеров.

На рис. 8 представлен результат работы классификации нейронной сети: зелёным цветом выделены траектории объектов, серым – траектории звёзд.

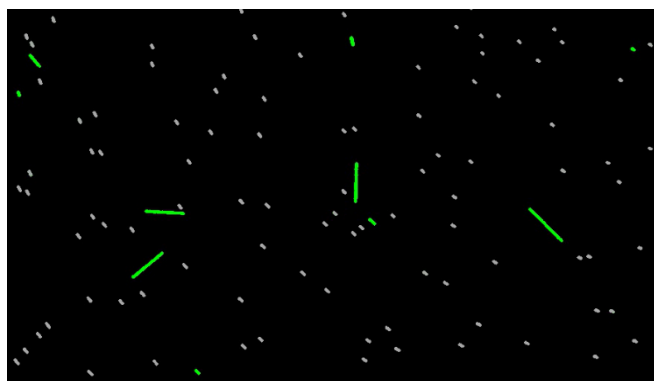


Рис. 8. Пример классификации траекторий

Модель показала следующие метрики на тестовой выборке:

ТАБЛИЦА II. МЕТРИКИ МОДЕЛИ U-NET

| Название классов | precision | recall | f1-score |
|------------------|-----------|--------|----------|
| 0 – Звёзды       | 0.74      | 0.98   | 0.84     |
| 1 – Объекты      | 0.61      | 0.97   | 0.84     |

#### IV. ЗАКЛЮЧЕНИЕ

В результате работы, на основе посчитанных метрик, было установлено, что SVM-классификатор превосходит нейросетевую модель U-Net по всем ключевым показателям.

Модель SVM демонстрирует практически идеальное качество по всем метрикам, что свидетельствует о высокой степени обнаружения и высокой степени точности классификации при наличии корректно извлечённых признаков. Однако, данный подход требует большого внимания к обработке входных данных человеком и плохо масштабируется на более сложные случаи.

Модель U-Net также демонстрирует высокую степень обнаружения объектов, но с классификацией пятен модель справляется хуже, что указывает на проблемы, связанные с дисбалансом классов и ограниченным объёмом обучающей выборки. Несмотря на это, способность нейросети выявлять объекты на минимально обработанных изображениях указывает на её потенциал в условиях реальных наблюдений, при дальнейшем увеличении объёма данных и применении современных нейросетевых архитектур.

Таким образом, SVM на текущем этапе представляет собой более точный инструмент для рассматриваемой задачи, тогда как нейросетевой подход требует дальнейшего развития, но обладает перспективами для применения в более универсальных и автоматизированных системах классификации объектов звездного неба.

#### СПИСОК ЛИТЕРАТУРЫ

- [1] Хайкин С. Нейронные сети: полный курс, 2-е издание. Москва: Издательский дом "Вильямс", 2006. 1104 с.
- [2] Вапник В.Н., Лернер А.Я. Распознавание образов при помощи обобщённых портретов // Автоматика и телемеханика. 1963. № 6. С. 774–780.
- [3] Cortes C., Vapnik V. Support-Vector Networks // Machine Learning. 1995. Vol. 20, No. 3. P. 273–297. DOI: 10.1007/BF00994018
- [4] Siji Joseph, Arun Pradeep - Object Tracking using HOG and SVM. Axis College of Engineering and Technology, 2017, 5 p.
- [5] Поляк М.Д. Модель стохастической искусственной нейронной сети в задачах распознавания образов // Научно-технические ведомости Санкт-Петербургского государственного политехнического университета. Информатика, телекоммуникации и управление. 2013. №1. С. 63–69.
- [6] Ronneberger O., Fischer P., Brox T. U-Net: Convolutional Networks for Biomedical Image Segmentation // In: Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015. Lecture Notes in Computer Science, vol. 9351. Cham: Springer, 2015. P. 234–241. DOI: 10.1007/978-3-319-24574-4\_28
- [7] Шквиро С.А., Поляк М.Д. Применение U-Net для выделения рядов виноградника по данным дистанционного зондирования Земли // V Международная конференция по нейронным сетям и нейротехнологиям (NeuroNT2024): Сб. докладов конференции, Санкт-Петербург, 20 июня 2024 года. Санкт-Петербург: СПбГЭТУ «ЛЭТИ», 2024. С. 106-109.
- [8] Николенко С., Кадурин А., Архангельская Е. Глубокое обучение Санкт-Петербург: Питер, 2018. 480 с.
- [9] A Robust Vision-based Algorithm for Detecting and Classifying Small Orbital Debris Using On-board Optical Cameras. Simulated environments generated using NASA Marshall Space Flight Center's Dynamic Star Field Simulator of on-board optical sensors and cameras. URL: <https://github.com/nnategh/AMOS2019> (дата обращения: 16.04.2025).
- [10] Shariar Hadji Hosseina, Marco Acernese, Tommaso Cardonaa, Giammarco Cialone, Federico Curianò, Lorenzo Mariani, Veronica Marini, Paolo Marzioli, Leonardo Parisi, Fabrizio Piergentili, Fabio Santoni. Sapienza Space debris Observatory Network (SSON): A high coverage infrastructure for space debris monitoring. // *Jornal of Space Safety Enginireeng*, 2020, Vol. 7, No. 1, 30-37 p.
- [11] Tiehan Lv, Burak Ozer, Wayne Wolf. A Real-Time Background Subtraction Method with Camera Motion Compensation. Multimedia and Expo, 2004. ICME '04. 2004 IEEE International Conference on. Vol. 1.
- [12] Zoran Zivkovic, Ferdinand van der Heijden. Efficient adaptive density estimation per image pixel for the task of background subtraction. *Pattern recognition letters*, 27(7):773–780, 2006.
- [13] Vedant Gaur. Lucas-Kanade Optical Flow Machine Learning Implementations. // *Jornal of Student Research*. 2022, Vol. 11, No. 3, 19 p.
- [14] Tsai Du.M., Chiu W.Ya., Lee M.H. Optical Flow-Motion History Image (OF-MHI) for Action Recognition. *Signal, Image and Video Processing*. 2014. Vol. 9, No. 8, 1897-1906 p.