Секция А. Астрофизика и физика Солнца. БШФФ-2026. С. 165–166.

УДК 004.8, 523.98
DOI: 10.62955/0135-3748-2026-165

АВТОМАТИЗИРОВАННАЯ РАЗМЕТКА ОБЪЕКТОВ НА ИЗОБРАЖЕНИЯХ СОЛНЦА, ПОЛУЧЕННЫХ СИБИРСКИМ РАДИОГЕЛИОГРАФОМ

Д.С. Чепель, В.В. Феденёв, С.А. Анфиногентов

Институт солнечно-земной физики СО РАН, Иркутск, Россия,
interceptor3001@gmail.com

AUTOMATED ANNOTATION OF SOLAR FEATURES ON IMAGES OBTAINED BY THE SIBERIAN RADIOHELIOGRAPH

D.S. Chepel, V.V. Fedenev, S.A. Anfinogentov

Institute of Solar-Terrestrial Physics SB RAS, Irkutsk, Russia, interceptor3001@gmail.com

Аннотация. Работа посвящена проблеме автоматизации разметки структур солнечной активности на радиоизображениях, получаемых Сибирским радиогелиографом (СРГ). Ежедневно архив наблюдений обсерватории пополняется большим количеством данных. Для проведения статистических исследований требуется составлять каталоги наблюдаемых на Солнце объектов. В статье описан методический подход к созданию конвейера для ручной и полуавтоматической разметки изображений для нужд машинного обучения. Представлен разработанный программный комплекс для составления каталогов структур солнечной активности: активных областей (АО), протуберанцев, корональных дыр и других. Обоснован выбор архитектуры сверточных нейронных сетей для решения дальнейших задач семантической и экземплярной сегментации.

Ключевые слова: Сибирский радиогелиограф, радиоизображения, машинное обучение, нейронные сети, сегментация изображений, разметка данных.

Abstract. This paper is devoted to the problem of automating the annotation of solar activity structures on radio images obtained by the Siberian Radioheliograph (SRH). Daily, the observatory's observation archive is replenished with a large amount of data. To conduct statistical research, it is necessary to compile catalogs of objects observed on the Sun. The article describes a methodological approach to creating a pipeline for manual and semi-automatic image annotation for machine learning needs. A developed software package for compiling catalogs of solar activity structures (active regions, prominences, coronal holes, etc.) is presented. The choice of the architecture of convolutional neural networks for solving further tasks of semantic and instance segmentation is justified.

Keywords: Siberian Radioheliograph, radio images, machine learning, neural networks, image segmentation, data annotation.

ВВЕДЕНИЕ

Сибирский радиогелиограф (СРГ) — один из наиболее передовых инструментов для исследования Солнца в микроволновом диапазоне, получающий радиоизображения на частотах 3–24 ГГц, обладающий высоким пространственным и временным разрешением (до 5 угл. сек и до 0.2 сек скважности) [Алтынцев и др., 2020]. СРГ ежедневно генерирует больше 100 Гб сырых данных. Поиск структур солнечной активности в большом массиве данных является долгим и трудоемким, в то время как для осуществления исследований в области физики Солнца требуются уже готовые каталоги. В связи с этим возникает острая необходимость разработки инструмента для их формирования. Данная работа направлена на создание полуавтоматического конвейера для разметки и сегментации объектов на радиокартах СРГ с задействованием методов машинного обучения.

РАЗРАБОТКА ИНСТРУМЕНТАРИЯ ДЛЯ РАЗМЕТКИ ДАННЫХ

Фундаментальной основой обучения любой нейросетевой модели выступает качественно размеченный набор данных. В связи со спецификой астрофизических данных (формат FITS, широкий динамический диапазон) стандартные программы разметки оказались малопригодны. Для решения этой задачи нами было спроектировано и разработано специализированное приложение Solar Annotator. Программа представляет собой автоматизированное рабочее место астрофизика-разметчика и обладает следующими возможностями.:

— Взаимодействие с вычислительным кластером: программа способна напрямую подгружать готовые (насчитанные) радиоизображения с вычислительного кластера Радиоастрономической обсерватории ИСЗФ СО РАН (Бадары).

— Выбор параметров излучения: для морфологического анализа структур в программе используются снимки интенсивности излучения (компонент Стокса *I*).

— Инструментарий полигональной сегментации: возможность точного ручного выделения границ сложных объектов (например протуберанцев или флоккулов) с помощью настраиваемых кистей и полигонов поверх радиокарт.

— Экспорт данных: наблюдатель может выделить нужные объекты на изображении и экспортировать их координаты и классы в формат JSON для последующей отправки в общую базу данных обсерватории.

На текущем этапе для формирования обучающей выборки используются изображения на

частоте 6 ГГц, которые рассчитываются каждые 15 мин в течение всего наблюдательного дня. Для удобства работы с таким временным рядом в интерфейсе программы реализован календарь, позволяющий пользователю легко выбирать и размечать данные за конкретные дни.

АРХИТЕКТУРА НЕЙРОСЕТЕВОЙ МОДЕЛИ

На начальном этапе работы с приложением вся разметка изображений производится пользователем исключительно вручную. Однако в дальнейшем планируется внедрение нейросети, которая будет выступать в роли ассистента: она будет предлагать предварительную разметку, помогая пользователю выделять объекты значительно быстрее и требуя от него лишь корректировки (подтверждения или исправления) результатов. Для решения задачи локализации солнечных объектов планируется применение архитектур класса U-Net [Ronneberger et al., 2015] или YOLO [Jocher et al., 2023]. Использование современных модификаций для экземплярной сегментации (Instance Segmentation, реализованной, например, в архитектуре YOLOv8-seg) позволяет не просто определить тип объекта, но и четко разделить независимые структуры одного класса (например, отделить одну активную область от другой, расположенной рядом).

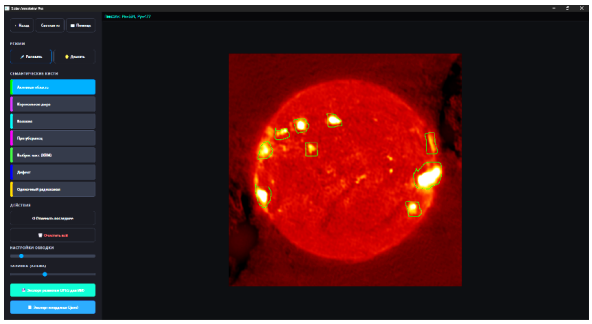
Процесс обработки будет включать предварительную нормализацию данных из формата FITS и логарифмическое масштабирование интенсивностей. На выходе модель будет генерировать вероятностные маски, где каждому пикселю исходного радиоизображения присваивается класс соответствующего физического объекта или фона. Далее с помощью классических алгоритмов машинного зрения полученные маски будут автоматически преобразовываться обратно в векторный формат JSON (в виде координат границ объектов) для бесшовной интеграции в общую базу данных.

ОЖИДАЕМЫЕ РЕЗУЛЬТАТЫ И ЗАКЛЮЧЕНИЕ

Создание размеченного набора данных с помощью Solar Annotator является важнейшим первым этапом на пути к автоматизации анализа данных СРГ. Внедрение обученных нейросетевых алгоритмов в систему обработки данных обсерватории ИСЗФ СО РАН позволит осуществить следующее.

1. Исключение рутинного ручного труда при составлении актуальных каталогов солнечных событий.
2. Стандартизация процесса выделения структур и снизить влияние субъективного человеческого фактора при разметке.
3. Значительное ускорение обработки ежедневного массива наблюдений.

В качестве подтверждения работоспособности созданного инструментария на рисунке показан интерфейс программы Solar Annotator с загруженным радиоизображением Солнца. Данный программный комплекс уже функционирует и помогает обсерватории в решении задач первичной обработки данных.



Интерфейс программы Solar Annotator с загруженным радиоизображением Солнца

СПИСОК ЛИТЕРАТУРЫ

Алтынцев А., Лесовой С., Глоба М., Губин А., Кочанов А., Гречнев В., Иванов Е., Кобец В., Мешалкина Н., Муратов А., Просовецкий Д., Мышьяков И., Уралов А., Федотова А. Многоволновой Сибирский Радиогелиограф // Солнечно-земная физика. 2020. Т. 6, № 2. С. 30–40.

Jocher G. et al. Ultralytics YOLOv8 [Электронный ресурс]. 2023. URL: https://github.com/ultralytics/ultralytics (дата обращения: 24.04.2026).

Ronneberger O., Fischer P., Brox T. U-Net: Convolutional Networks for Biomedical Image Segmentation // Medical Image Computing and Computer-Assisted Intervention (MICCAI). 2015. P. 234–241.