

Отчет по курсовой работе за неделю

Дата: 15.10.2020

Научные руководители: Герасимов С.В., Мещеряков А.В.

Студент: Немешаева Алиса

Курс: 4

1. На этой неделе продолжалось усовершенствование алгоритма детекции: теперь вместо того, чтобы сканировать каждый патч отдельно и детектировать на нем скопления, алгоритм собирает все патчи в общую карту для большого сегмента неба, чтобы детектировать скопления на нём. Это помогает уменьшить количество false positives, которые появлялись из-за несовершенства старого алгоритма.
2. Были добавлены новые параметры сканирования: теперь для каждого детектированного скопления известны такие свойства, как:
 - `min_rad`, `mean_rad`, `max_rad` - минимальный, средний и максимальный радиусы предсказанной маски скопления.
 - `min_pred`, `max_pred` - минимальный и максимальный prediction index (значение предсказанной маски) скопления.
3. В задаче детекции скоплений важной проблемой является определение массы скопления. Поэтому бывает полезно искать корреляции параметров детектированных скоплений из известных каталогов с массой.

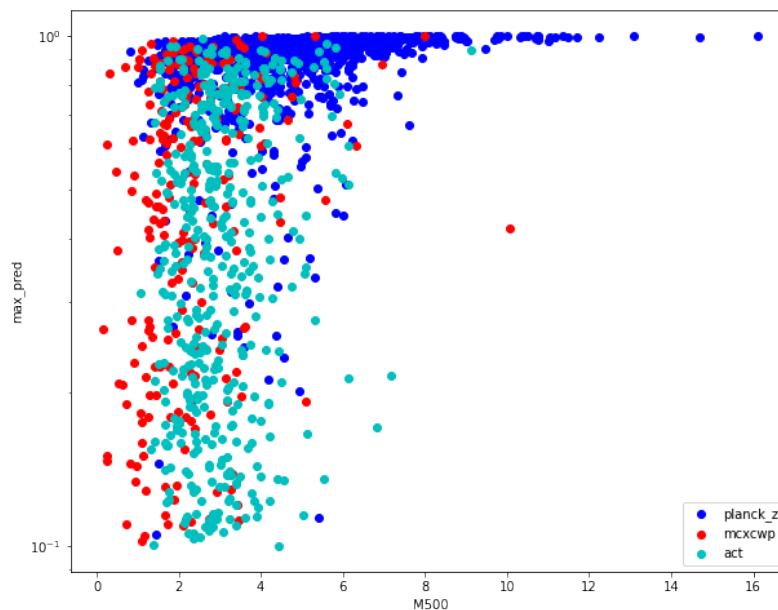


Рис. 1: График соответствия M500 и max_pred для скоплений true positives.

4. Кроме того, имеет смысл проверить соответствие параметров z (красное смещение) и M500 (масса) для скоплений true positives для того, чтобы иметь возможность формировать каталоги для обучения других моделей - скопления с некоторыми комбинациями параметров скорее всего не получится сегментировать в данных Planck.

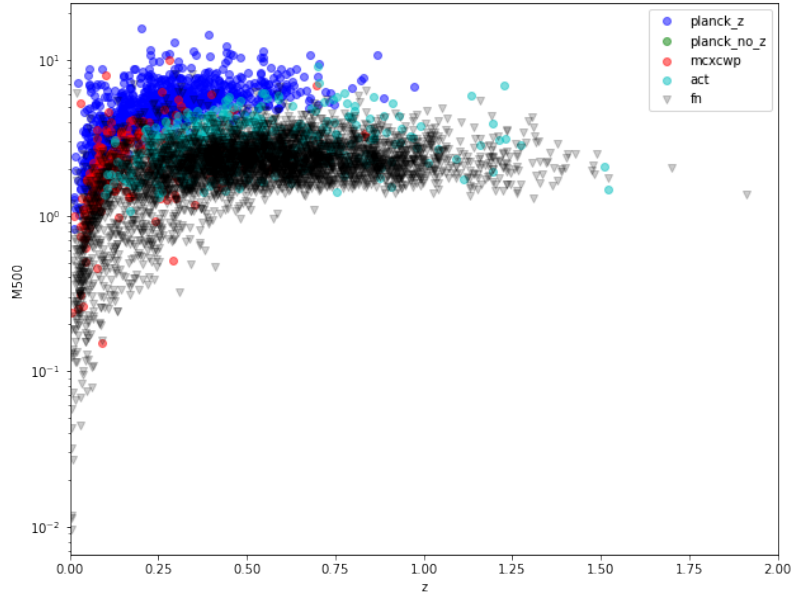


Рис. 2: График соответствия M500 и z для скоплений true positives и false negatives.

- Главной целью было создание окончательного каталога с наименьшим количеством false positives. Для этого сравнивались каталоги для разных значений thr. Кроме того, теперь ещё и исследовался вопрос выбора шага сканирования - с уменьшением шага recall почти не меняется, но количество fp уменьшается, что скорее всего является признаком более чистого каталога.

	planck_z	planck_no_z	mcxcwp	fp	all	step	thr
0.1	0.875	0.666667	0.04	134	242	64	0.1
0.1	0.925	0.722222	0.08	136	244	32	0.1
0.1	0.925	0.833333	0.08	145	253	16	0.1
0.1	0.950	0.777778	0.10	119	227	8	0.1

Рис. 3: Таблица сравнения recall и общего количества найденных объектов в тестовой области для разных значений шага

- Прямо сейчас идёт создание каталогов для thr = 0.1 и 0.2 и step = 8 для всего неба.
- Результаты текущей работы были продемонстрированы на семинаре группы ИКИ РАН по исследованию данных eRosita (презентация прилагается).

Общее количество строк кода за эту неделю: 202