

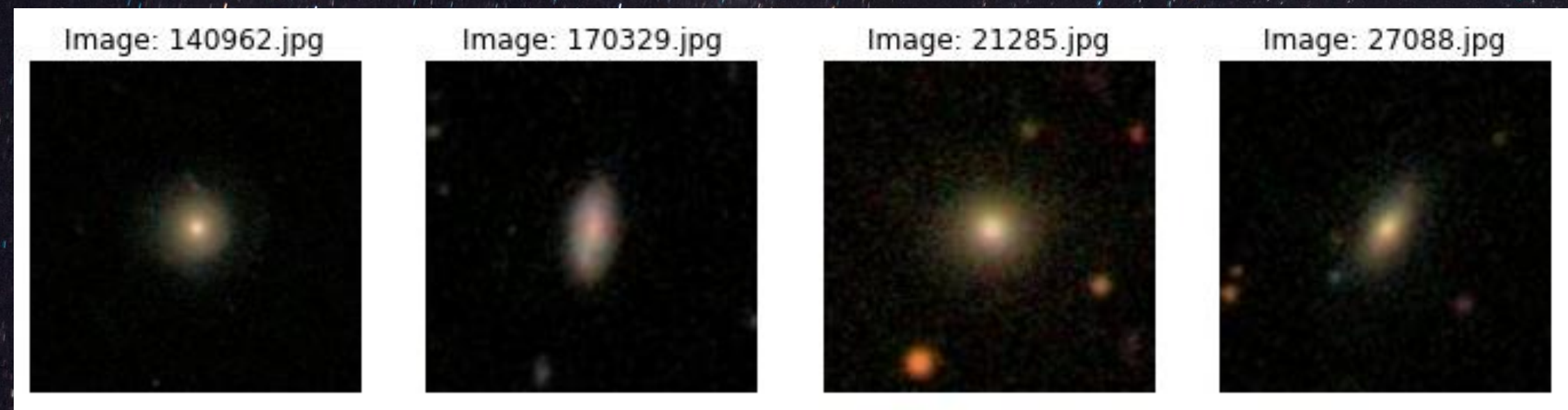
NARZĘDZIE DLA ASTRONOMÓW?

GABRIELA MAJSTRAK

MICHAŁ KUKLA

CEL PROJEKTU:

STWORZENIE MODELU DO KLASTERYZACJI ZDJĘĆ GALATYK

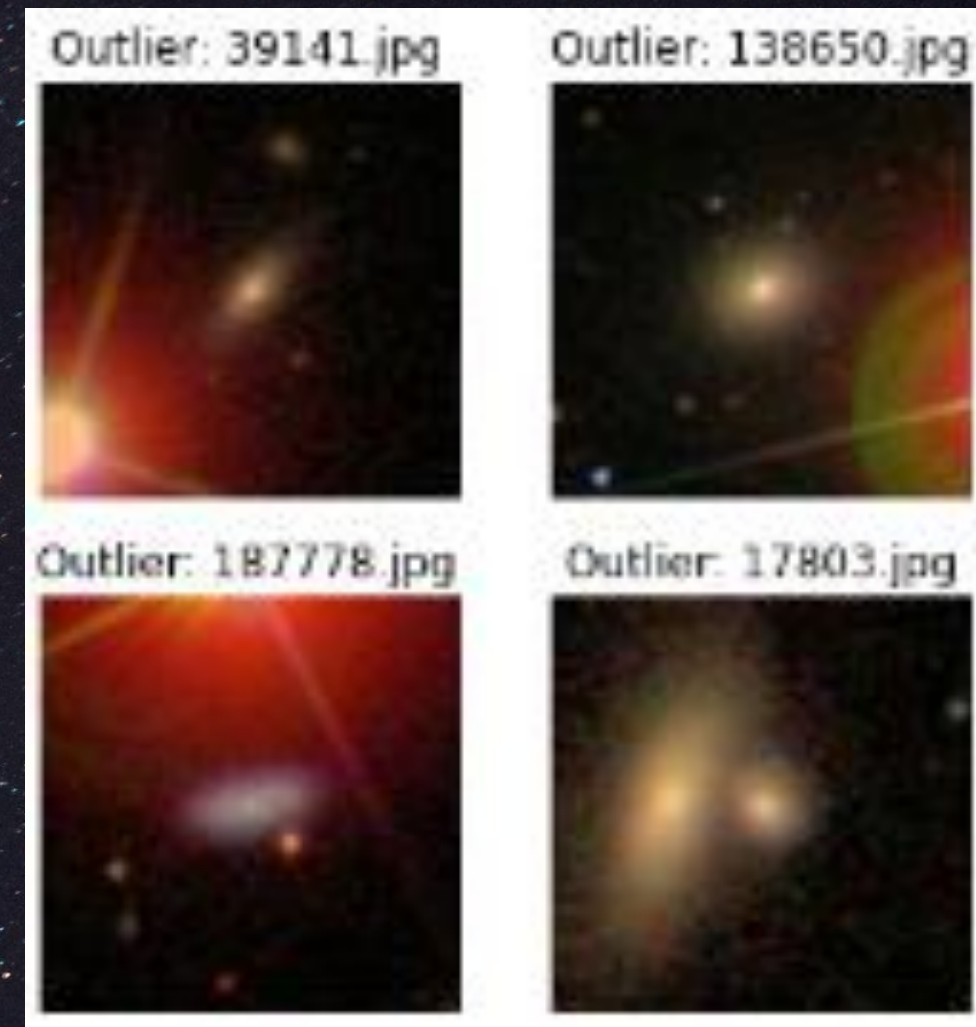


przykładowe zdjęcia, którymi
dysponowaliśmy

ed

Nasz projekt rozpoczęliśmy od przeprowadzenia eksploracyjnej analizy danych i dzięki niej wyciągnęliśmy następujące wnioski:

- wszystkie zdjęcia są tego samego rozmiaru
- występują zdjęcia, które znacznie odbiegają kolorystycznie od reszty
- kluczowe elementy zdjęć znajdują się w ich centrum



FEATURE ENGINEERING

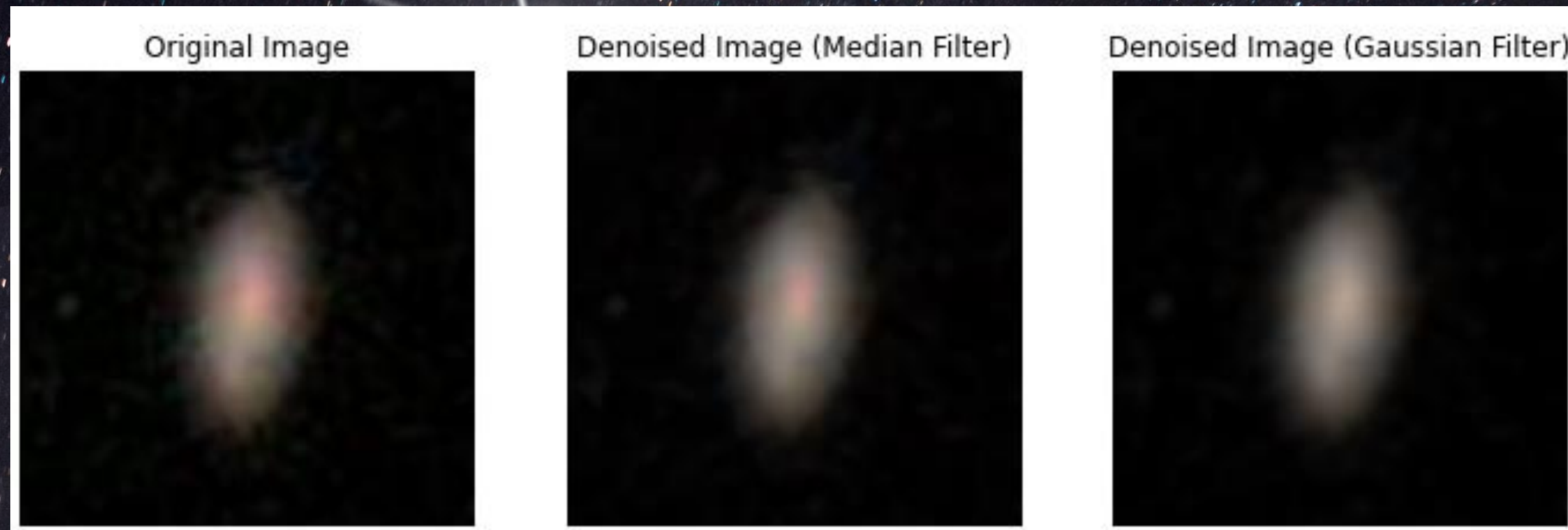
Ze względu na to, że galatyki, czyli to co nas najbardziej interesuje znajdują się w centralnej części większości zdjęć co sprawdziliśmy na etapie EDA, postanowiliśmy nasze zdjęcia 128x128 obciąć z każdej strony o 25 pikseli

Sample Cropped Image



USUWANIE

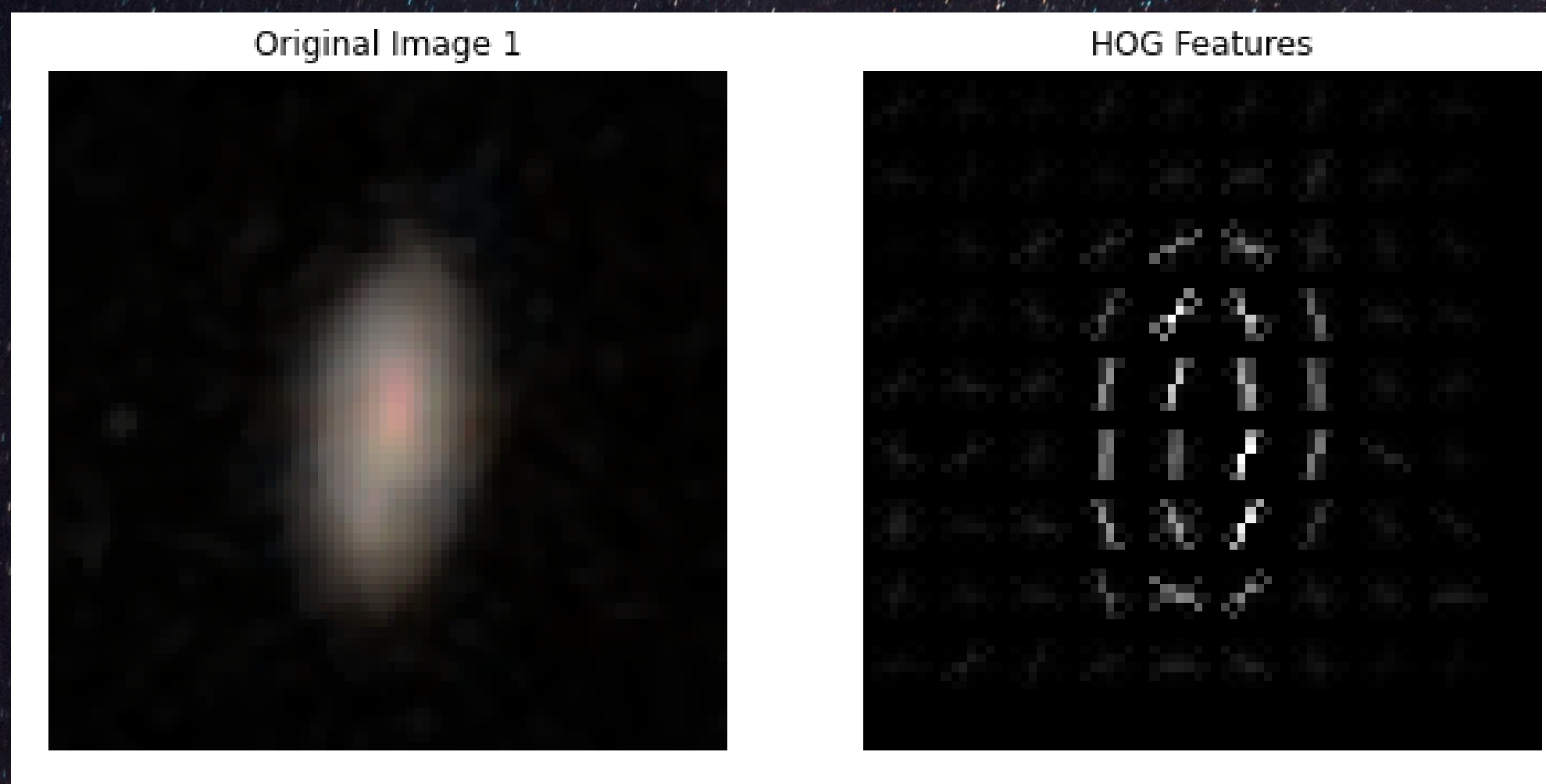
Szumów
Próbowaliśmy to zrobić. Usuwać szumy ze zdjęć,
wykorzystaliśmy w tym celu filtr medianowy i
gaussowski



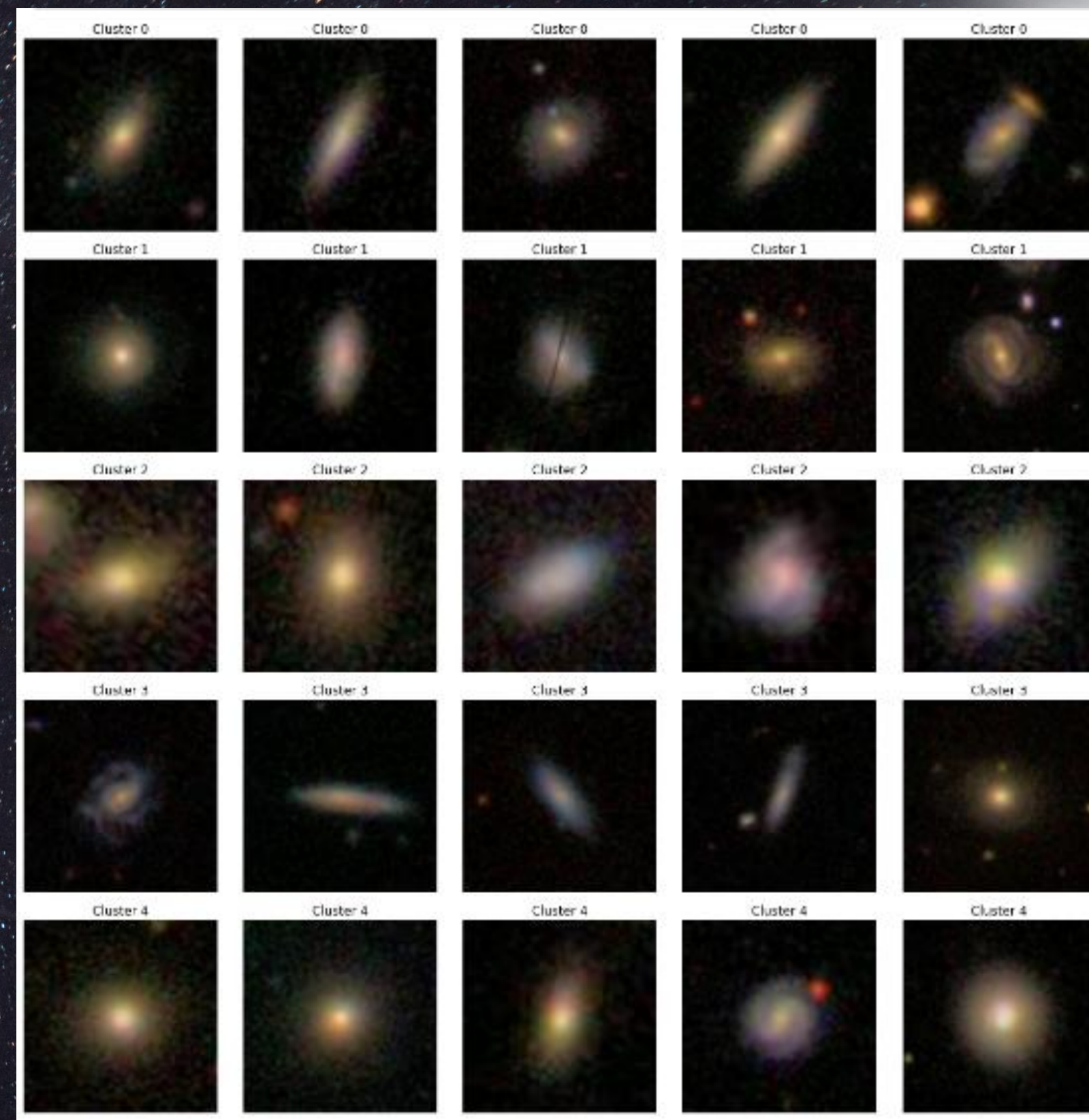
Więcej szczegółów zachowuje filtr medianowy, więc do
daleszej pracy zosatwiliśmy zdjęcia odszumione właśnie
tym filtrem

HISTOGRAM ORIENTED GRADIENTS (HOG)

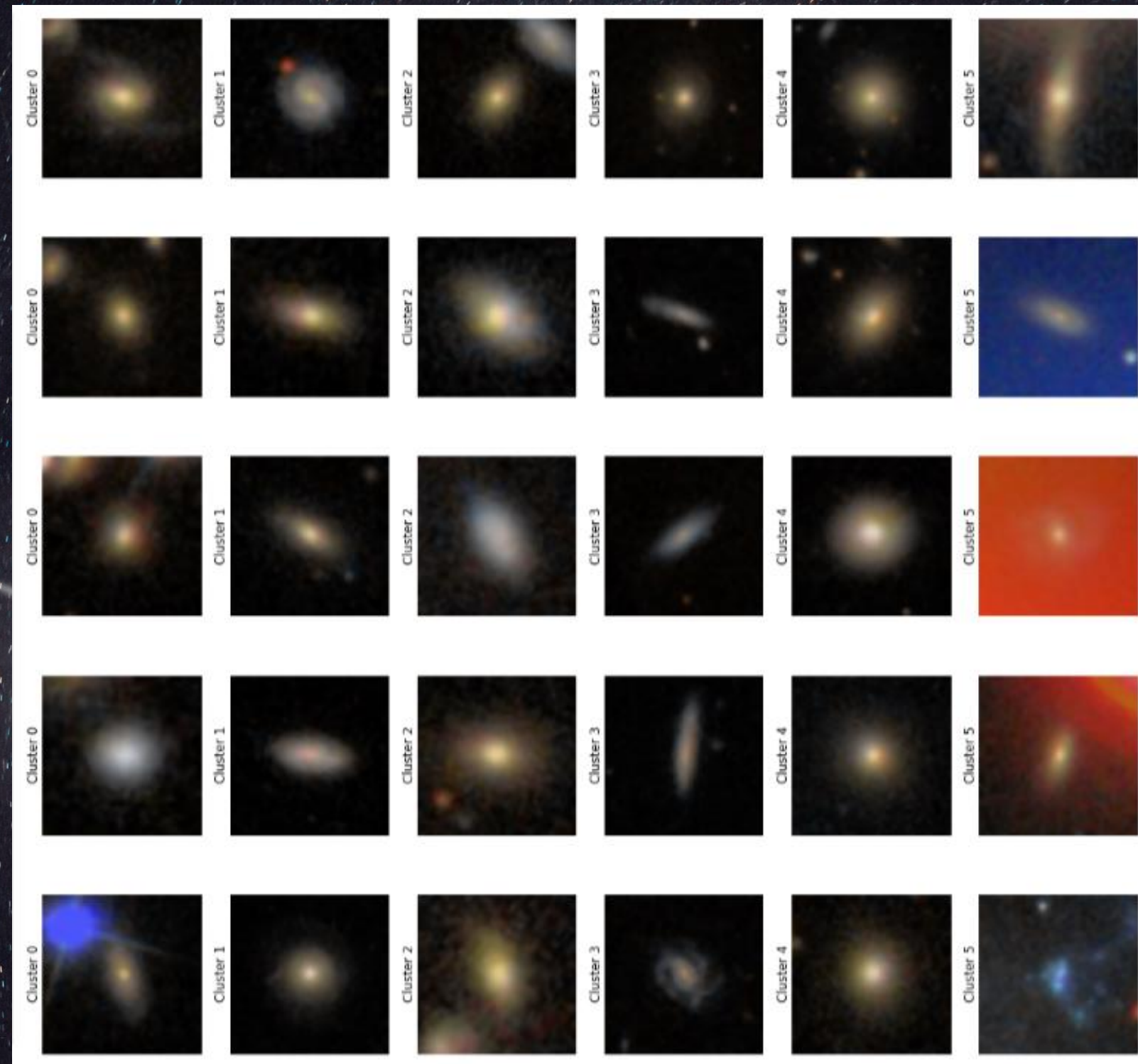
Dla zdjęć odszumionych filtrem Medianowym zastosowaliśmy algorytm ekstrakcji cech Hog, później tworzone modele testowaliśmy również na dla zdjęć na których został zastosowany ten algorytm



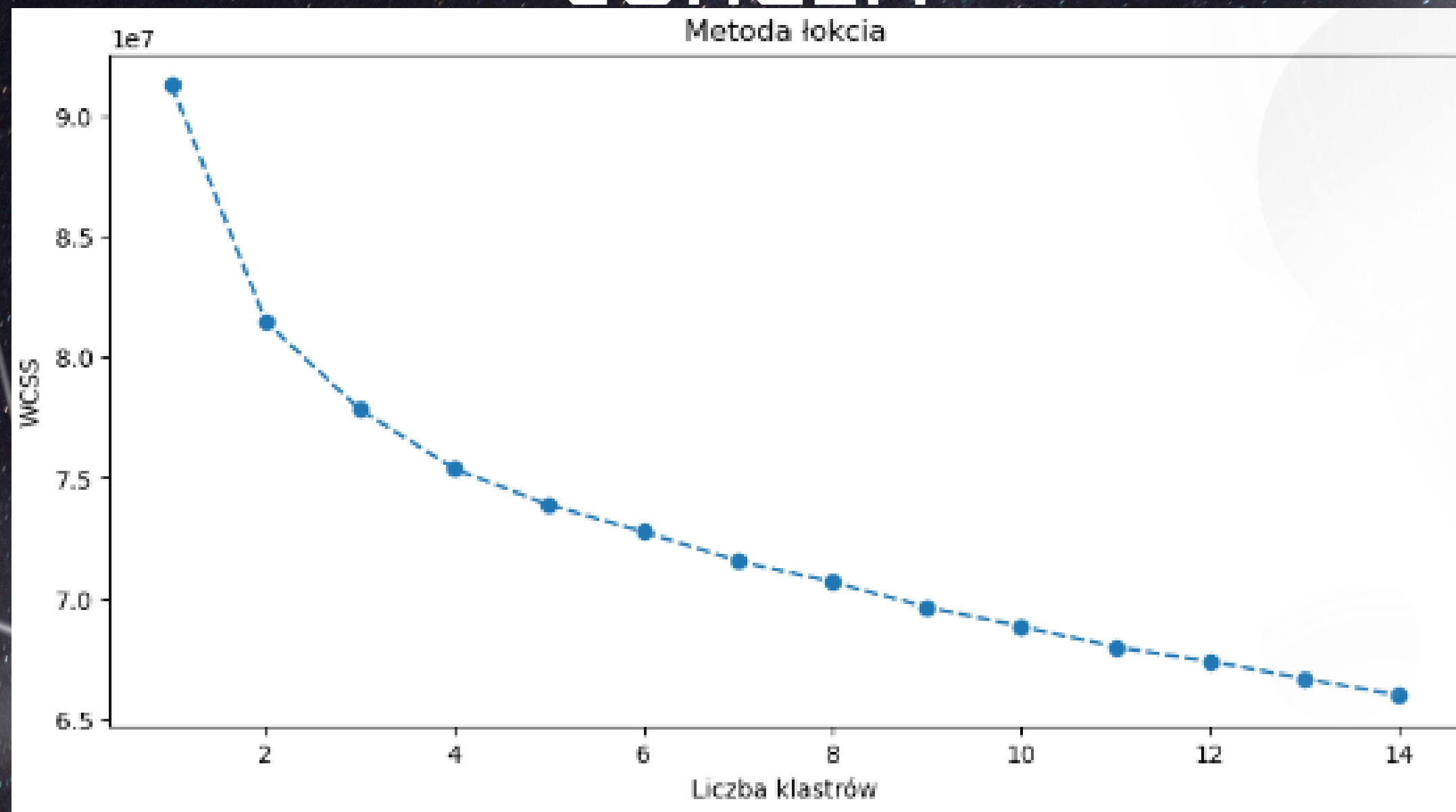
WYNIKI WSTĘPNEJ KLASTERYZACJI NA ZBIORZE OBCIĘTYCH ZDJĘĆ



WYNIK WSTĘPNEJ KLASTERYZACJI NA ZBIORZE ZDJĘĆ ODSZUMIONYCH FILTREM MEDIANOWYM

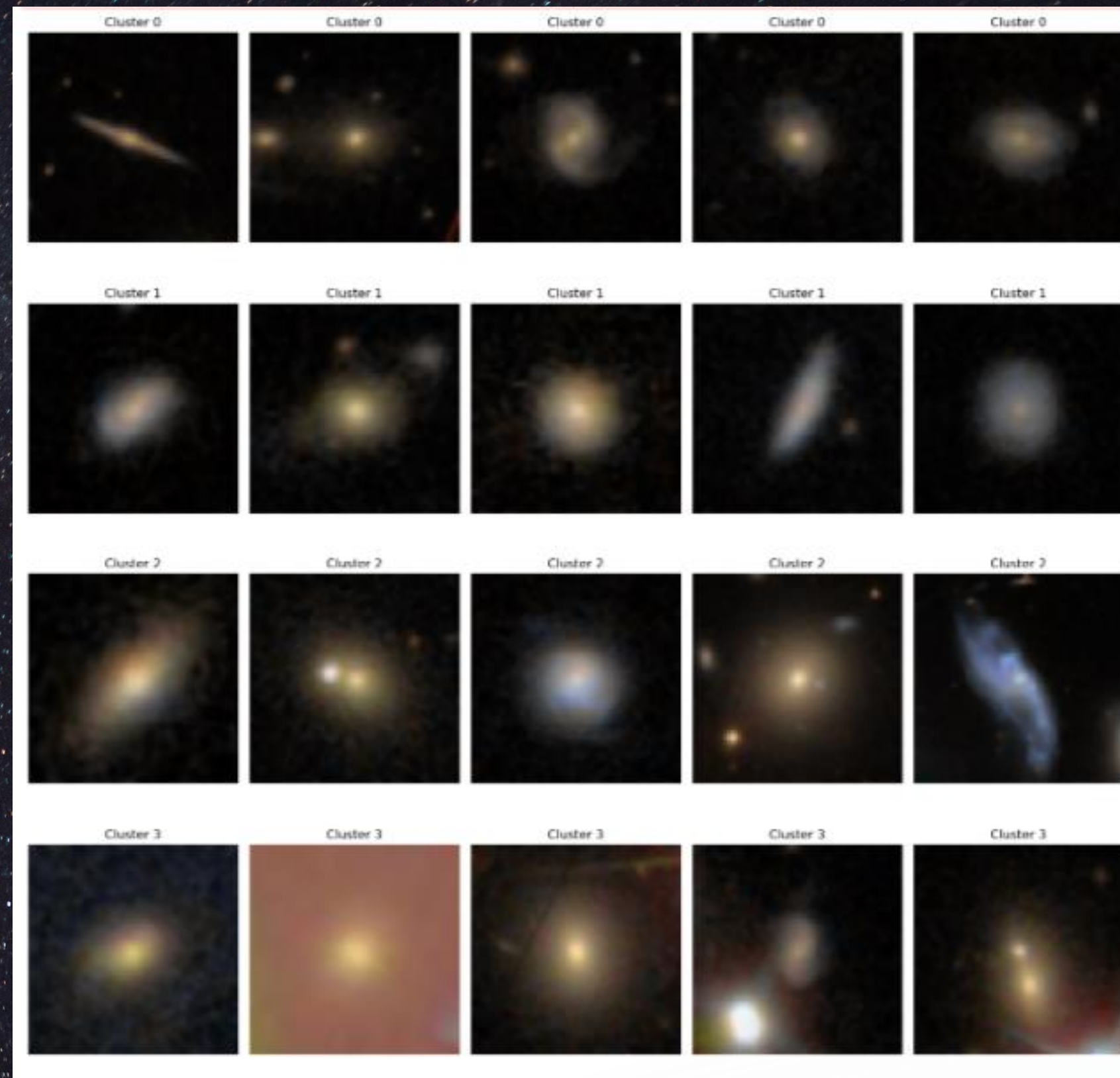


METODA ŁOKCIA



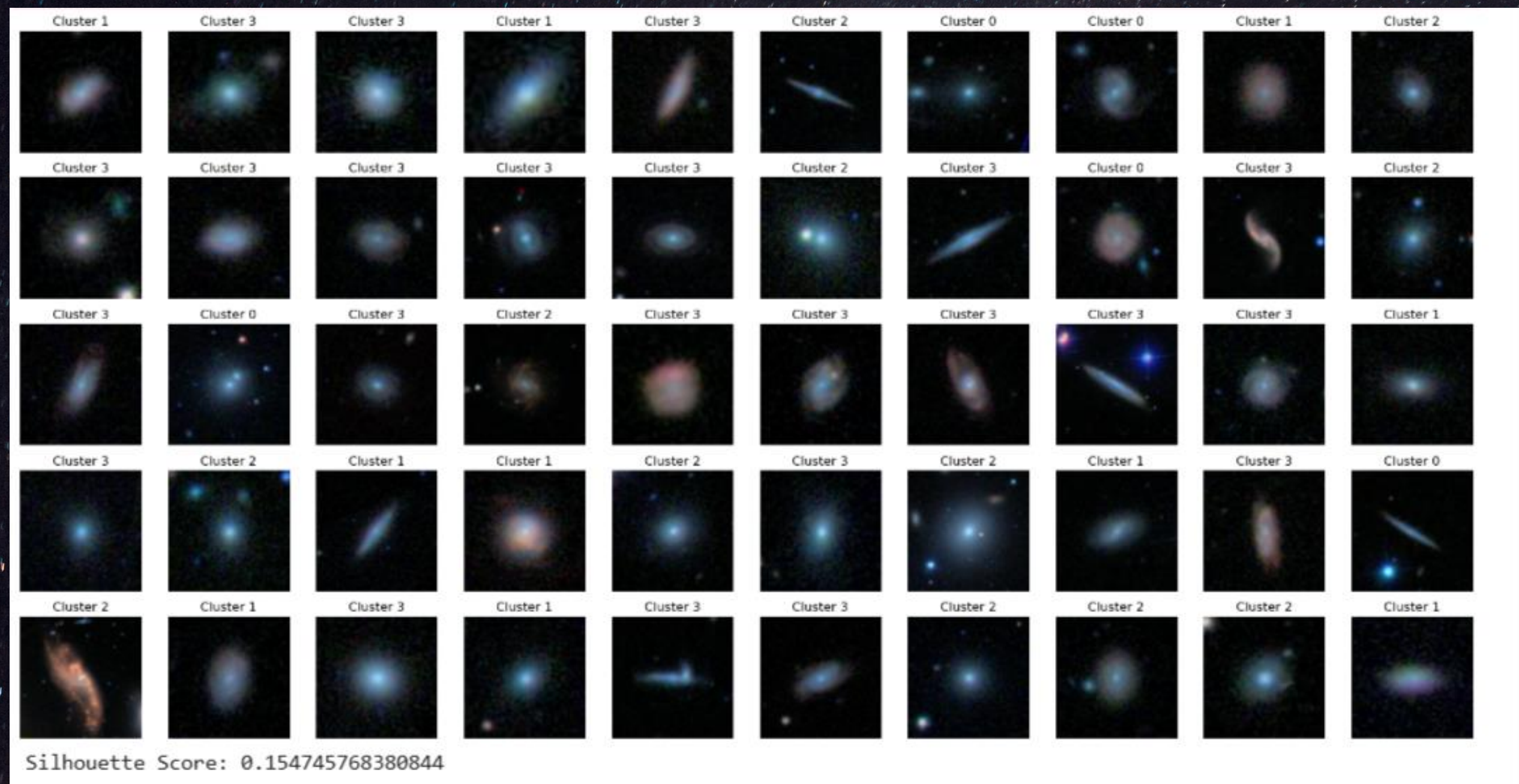
MODEL DLA 4 KLASTRÓW

Patrząc na wykres z poprzedniego slajdu postanowiliśmy przetestować model dla 4 klastrów. Jednak wyniki nie były zadowalające.



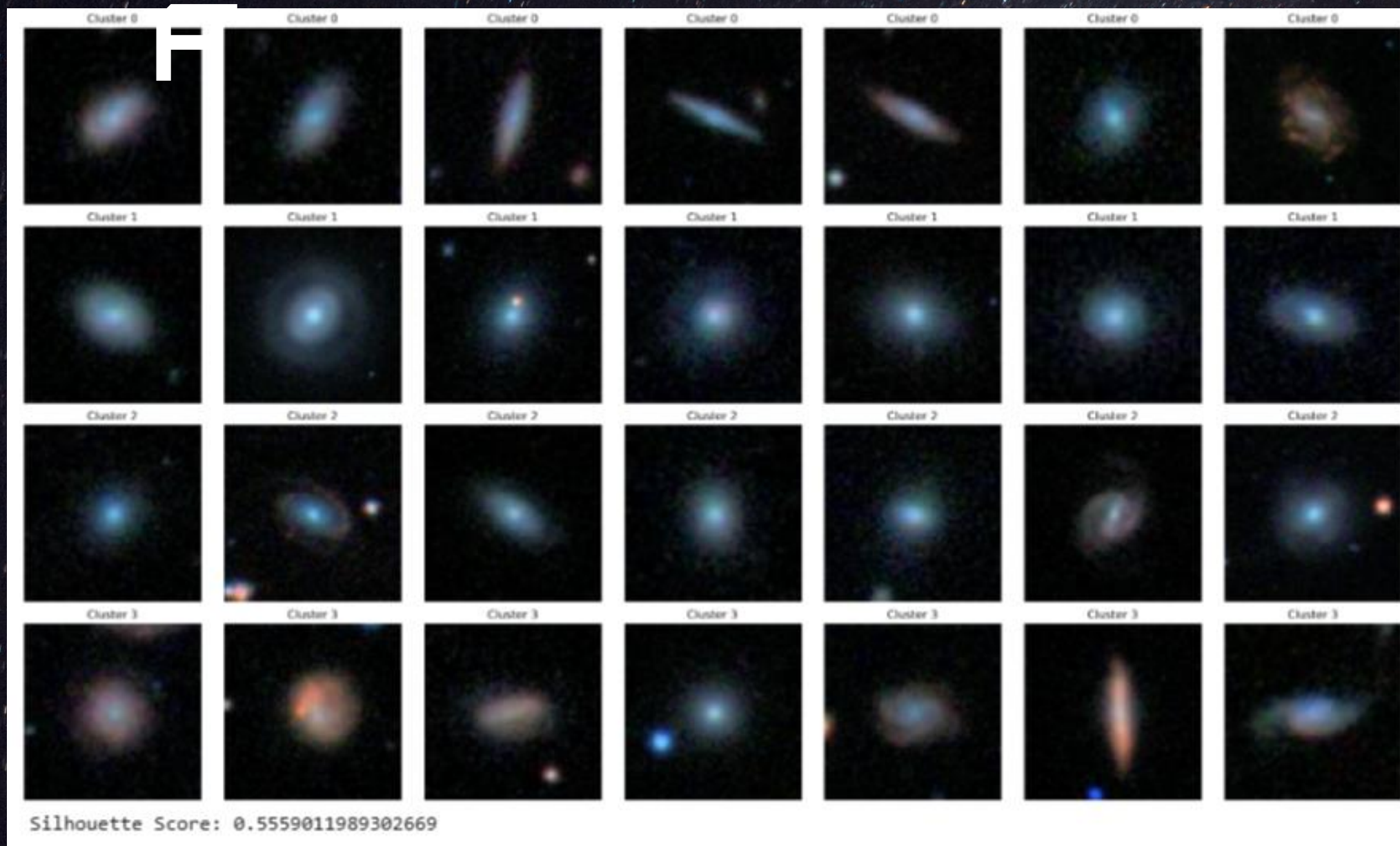
EKSTRAKcja CECH SIFT i SURF

Wykrywanie punktów charakterystycznych
Orientacja dla każdego punktu
charakterystycznego na podstawie lokalnych
gradientów obrazu



niezadowalający
Silhouette
score
i
niezadowalające
klastry

SUR



Widać, że SURF radzi
sobie już dużo lepiej

EKSTRAKCJA CECH

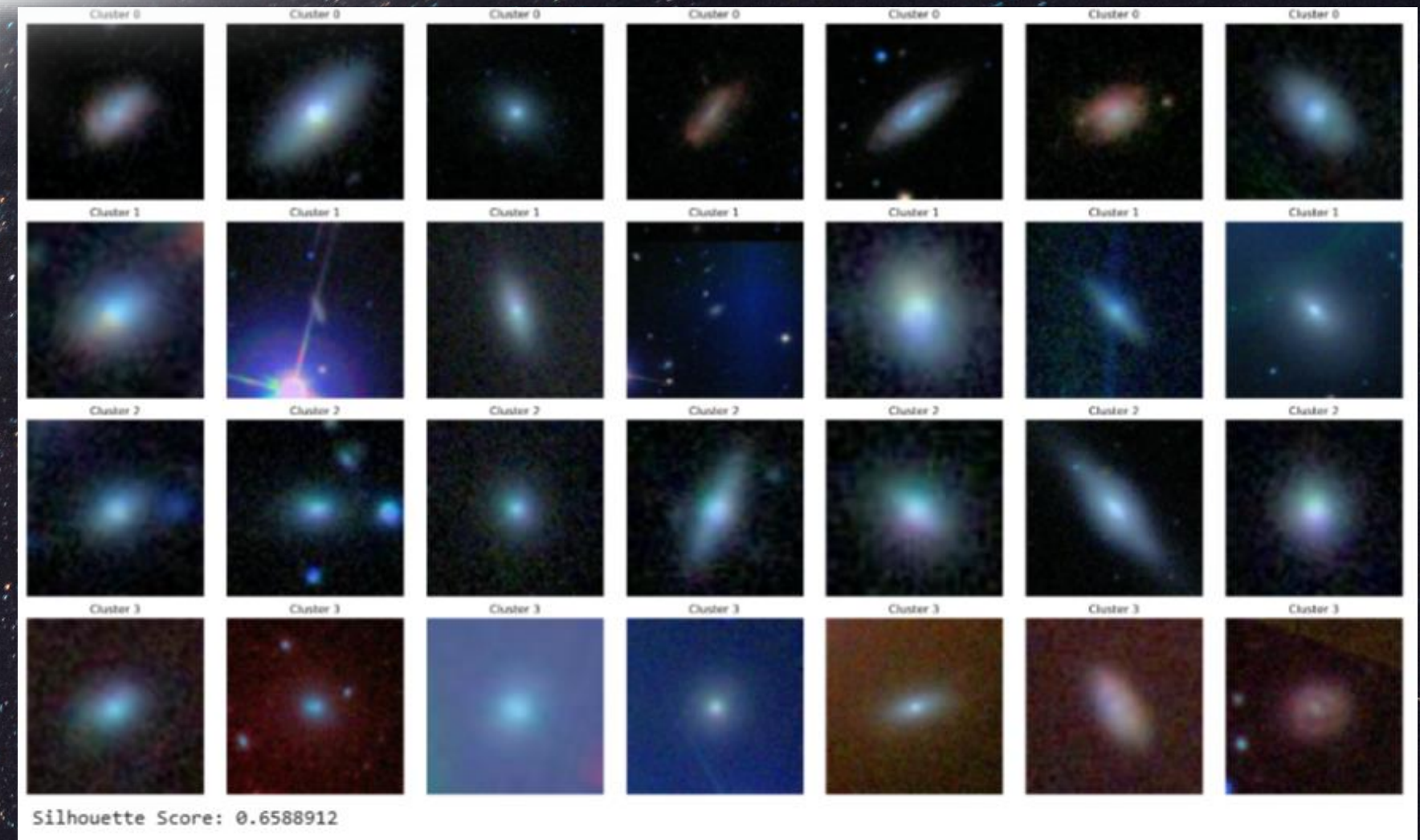
HISTOGRAMY

KOLORÓW

ODSZUMIONE

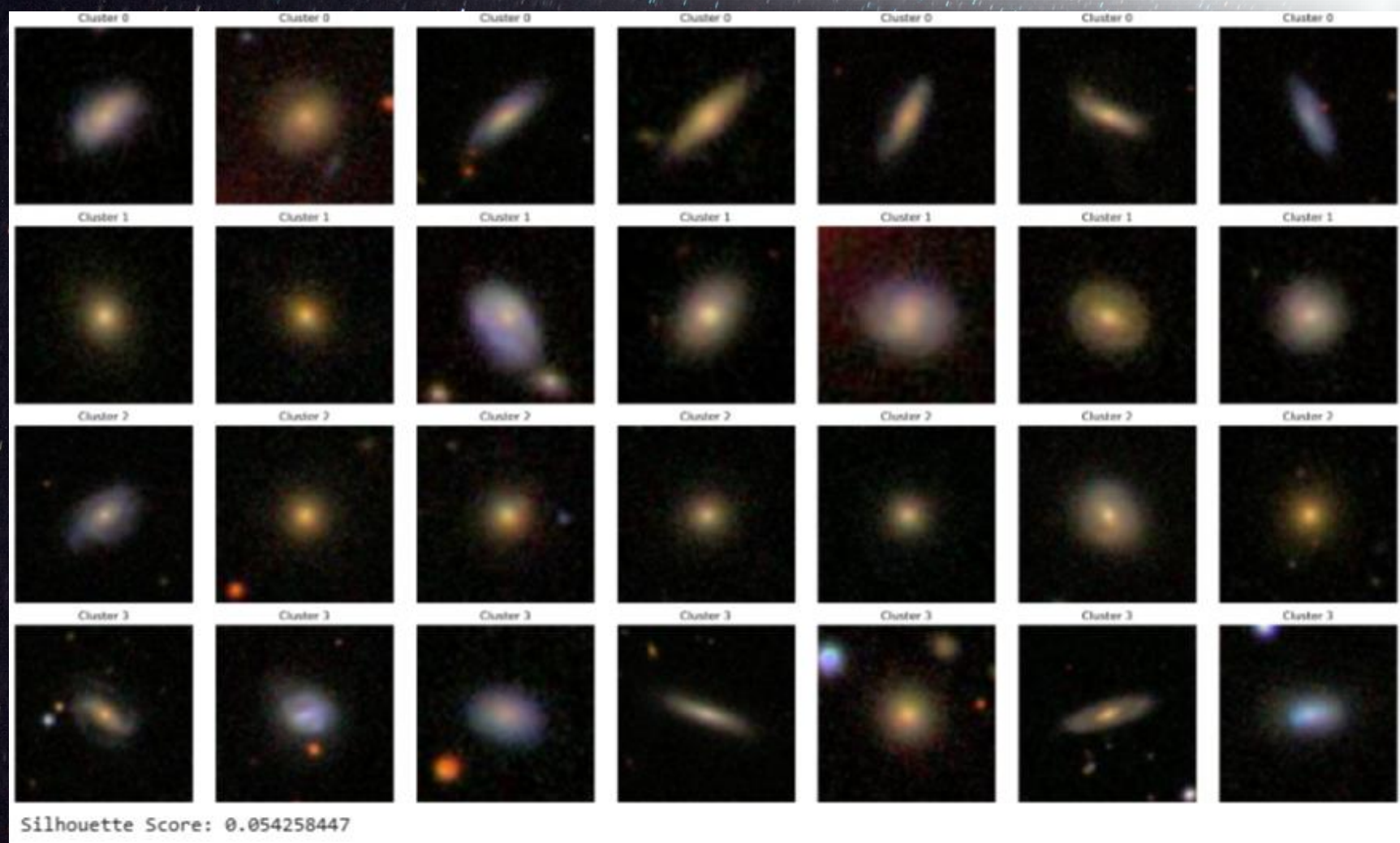


cropped

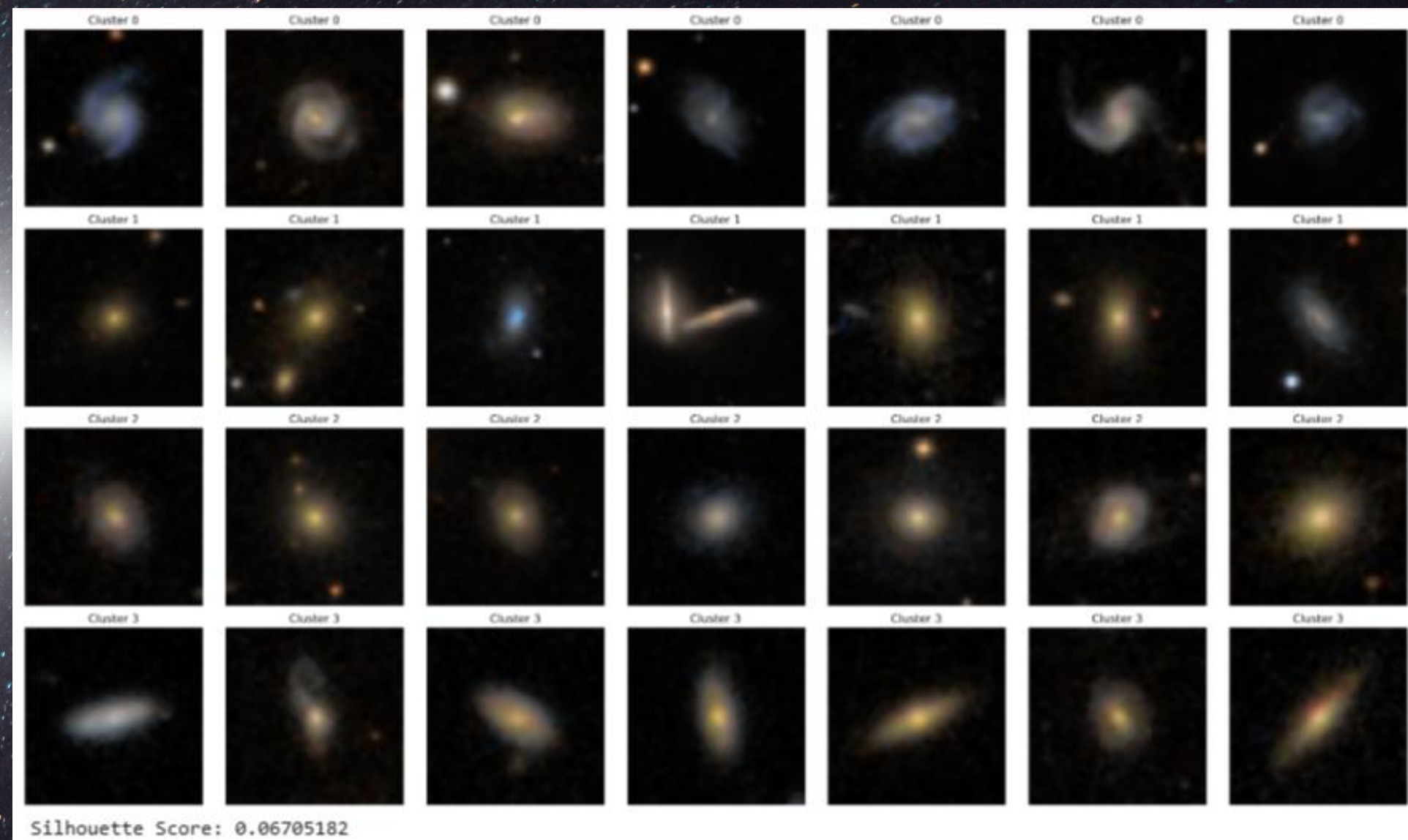


DEEP LEARNING embeddings

cropped

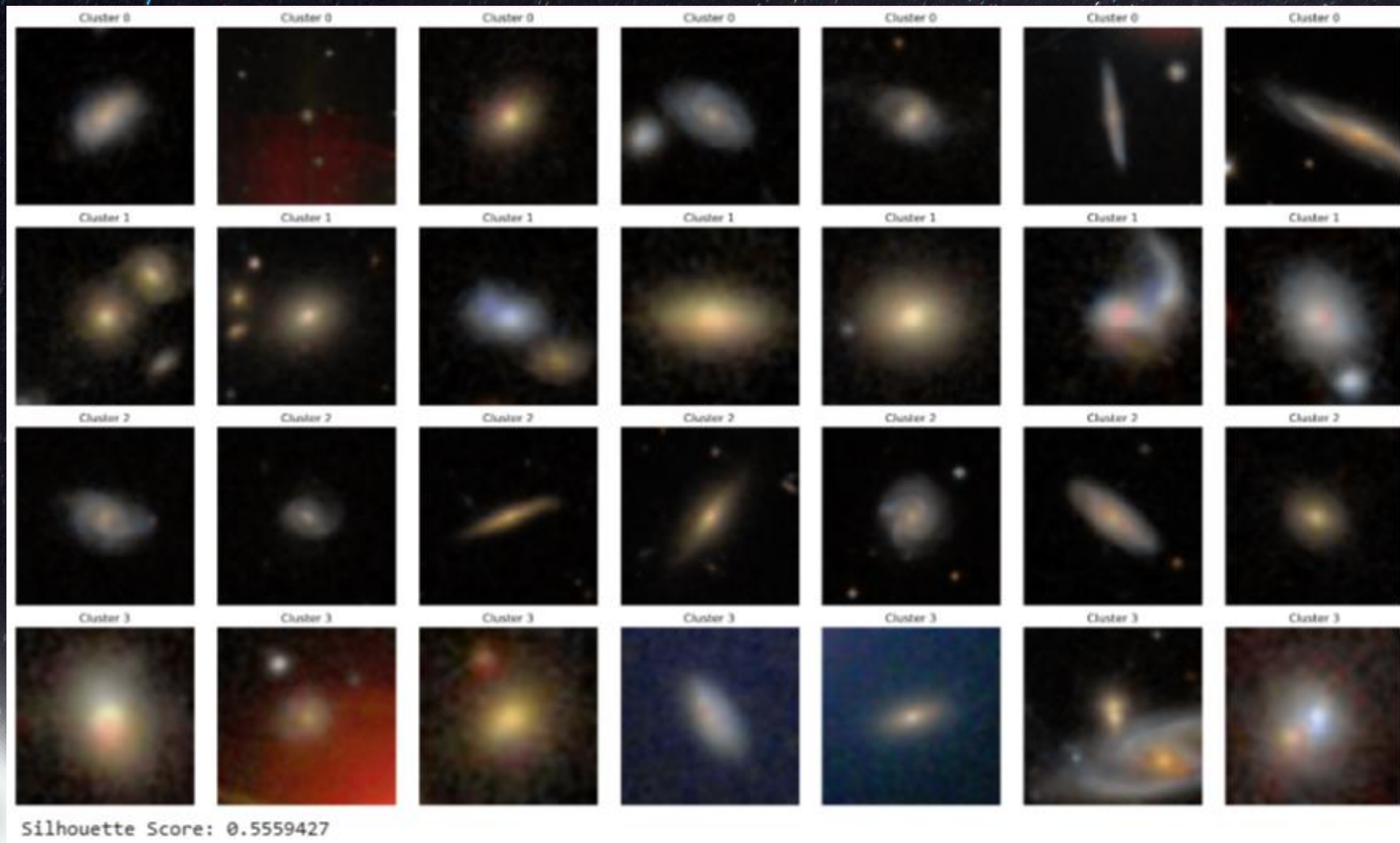


denoised



słaby Silhouette Score, ale wydaje się jakby w pierwszym klastrze były galaktyki spiralne

AUTOENKODERY



WNIOSKI

- Najlepsze Slithouette Score otrzymaliśmy dla histogramów kolorów i klastry wydają się rzeczywiście sensowne
- Oprócz tego patrząc po otrzymanych klastrach sensowny wydaje się też surf
- Słaby Slithouette Score dla deep learning embeddings, ale model wydaje się klasteryzować galatyki spiralne



DZIĘKUJEMY ZA
UWAGĘ