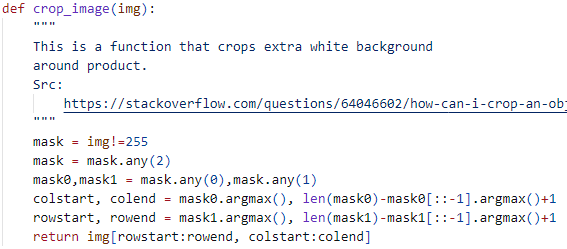
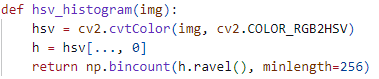
**BÀI THỰC HÀNH SỐ 4**

Sử dụng bộ dữ liệu có sẵn Digikala (Tên một hệ thống siêu thị lớn ở khu vực châu Á). Bộ dữ liệu này bao gồm các bức ảnh chụp các mặt hàng trên trang web với các màu sắc khác nhau. Trong bài thực hành này sinh viên sử dụng lược đồ histogram của kênh màu H trong HSV để làm vector nhận dạng. Hãy trích xuất đặc trưng cho tập ảnh, với mỗi ảnh thu được một vector histogram (H) và sử dụng nó để nhận dạng.

* Hàm crop\_image để lược bỏ những vùng trắng xung quanh ảnh, tăng độ chính xác khi áp dụng thuật toán.



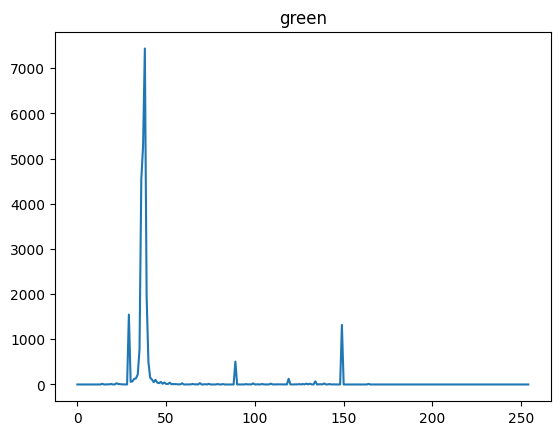
* Hàm hsv\_histogram để tính histogram của hình ảnh.



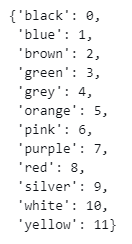
* Lấy tất cả các đường dẫn ảnh:



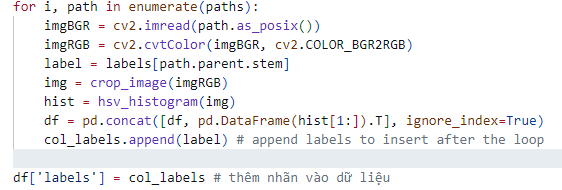
* Vẽ biểu đồ histogram của ảnh ngẫu nhiên:



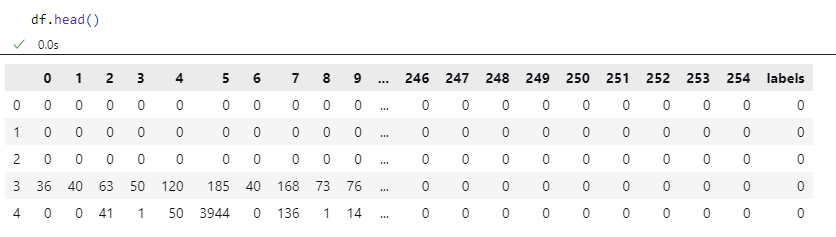
* Gán nhãn cho từng màu:



* Tính vector histogram cho từng ảnh:



* Dữ liệu thu được:



* Chia dữ liệu thành 2 tập train và test với tỷ lệ 80/20.
* Áp dụng thuật toán KNN và SVM thu được độ chính xác như sau:

