1. Quanv filter có thể có số lượng tham số ít hơn conv với cùng kích thước

2. Vì input và output của quanv giống nhau nên có thể tận dụng chỗ này để tối ưu.

Ví dụ 2x2 quanv filter trượt qua ảnh 28 x 28

Đầu ra sẽ là 14 x 14 x 4

4x4 quanv filter trượt qua ảnh 28 x 28

Đầu ra sẽ là 7 x 7 x 16

Khác với conv filter

1 2x2 conv filter trượt qua ảnh 28 x 28 đầu ra sẽ là 14 x 14 x 1

Phải 4 2x2 conv filter thì mới đạt được giống 1 quanv filter

=> tốn ít tham số hơn

Tổng quát, nếu cần output có deep là x, thì dùng math.ceil [ deep / kernel\_size\*\*2 ] quanv filter là được.

=> time consume nhảy bậc => hỏi thầy độ phức tạp này là tuyến tính p k?