

Przetwarzanie i Rozpoznawanie Obrazów

Deep Learning: Rozpoznawanie dróg na zdjęciach satelitarnych

1. Wstęp

W ramach laboratorium przedmiotu Przetwarzanie i Rozpoznawanie Obrazów przygotowano i przetestowano konwolucyjną sieć neuronową, która ma za zadanie oznaczać drogi na dostarczonym zdjęciu satelitarnym.

2. Opis podejścia pierwotnego

Postanowiliśmy do wykonania zadania użyć wykorzystywanej na laboratorium biblioteki keras. Wejściem utworzonej sieci jest fragment obrazka 5x5 (+3 kanały kolorów), natomiast na wyjściu znajdują się dwa neurony, z których każdy zwraca prawdopodobieństwo przyporządkowania do odpowiedniej klasy (droga, nie droga). Sieć składa się z pięciu warstw, kolejno: warstwa konwolucyjna składająca się z czterech filtrów 3x3, warstwa MaxPooling2D, warstwa Dropout, warstwa Flatten i na końcu warstwa Dense. Sieć dla każdego piksela obrazu z osobna (po kolei) zwraca informację, czy jest tam droga, czy też nie. Otrzymany wynik dla wszystkich ponad dwóch milionów pikseli (rozmiar dostarczonych danych to zdjęcia 1500x1500), jest następnie z powrotem łączony w tablicę i zwracany na wyjściu. Otrzymane obrazki można wyświetlić i porównać zarówno z dostarczoną etykietą, jak i z właściwym zdjęciem. Tak utworzona sieć w pierwszym etapie uczenia otrzymała na wejściu 100 obrazów uczących (dla wszystkich pikseli na nich występujących). Niestety, jak się potem okazało, wyniki były na tym etapie dalekie od satysfakcjonujących.

3. Opis podejścia drugiego

Pierwotna sieć przy początkowym sposobie uczenia niestety uznała za najkorzystniejsze przydzielanie klasy *brak drogi* dla każdego piksela, co mimo dużej teoretycznej skuteczności wynikającej z ogromnej przewagi liczbowej pikseli tej klasy nie było satysfakcjonującym rozwiązaniem. Jako że zdecydowana większość pikseli obrazka nie należy do żadnej drogi, nauczona sieć była dużo bardziej skłonna do przydzielania przypadków testowych do klasy większościowej. Postanowiliśmy skorzystać z techniki oversamplingu poprzez wielokrotne skopiowanie przykładów opisujących drogi (podania ich wielokrotnie na wejście naszej sieci neuronowej). Na tym etapie zdecydowaliśmy, że każdego z tych fragmentów użyjemy do uczenia 11 razy. Uczenie zostało tym razem przeprowadzone dla 50 zdjęć uczących. Otrzymane

wyniki były znacznie lepsze, sieć w poprawny sposób rozpoznawała większość dróg, pojawiły się też szumy (pojedyncze piksele lub grupy pikseli wskazane jako drogi, które drogami nie były). Ich znaczenie i wielkość zmniejszyliśmy korzystając z filtra medianowego o rozmiarze 5x5. Przykładowy wynik zaprezentowano na rysunku 1.



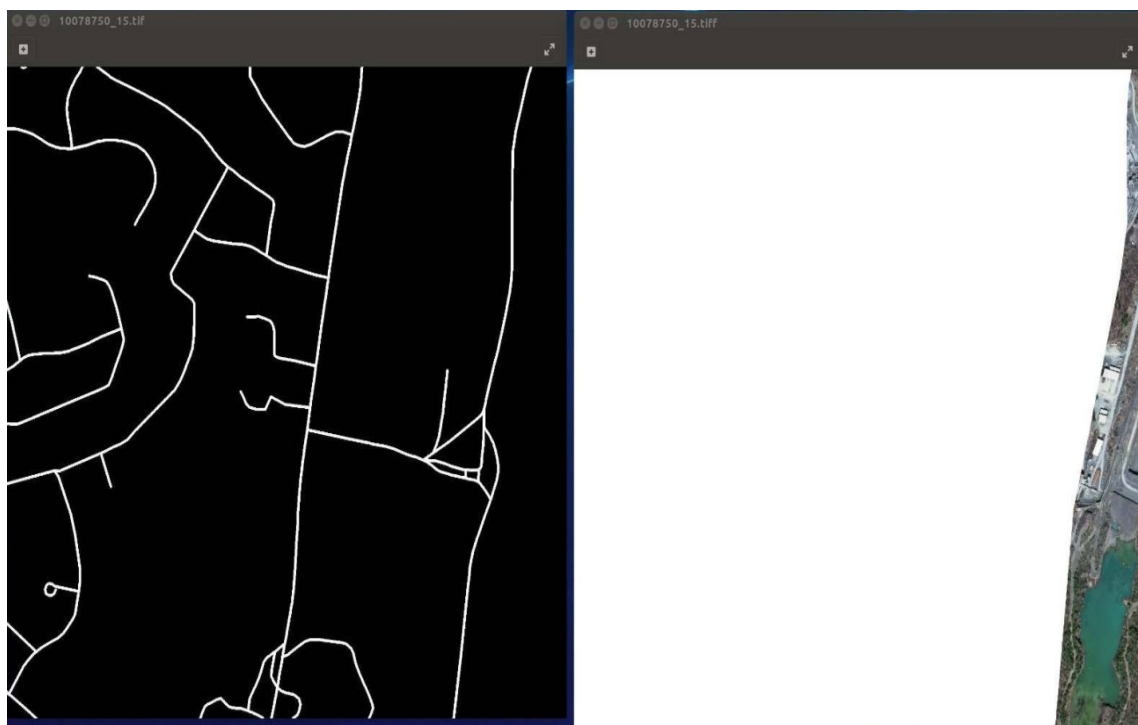
rys. 1. Wynik działania sieci po pierwszym etapie uczenia.

4. Opis prób polepszenia podejścia drugiego

Jako że sieć w drugim podejściu uznawała też wiele domów i innych obiektów za drogi, postanowiliśmy w dalszym procesie uczenia przetestować douczenie sieci dla innych licznosci powielonych przykładów. Jako że w pierwszej iteracji przykłady pozytywne były powielane jedenastokrotnie, przetestowaliśmy douczenie sieci od tego momentu dla kilku wartości z zakresu od czterech do szesnastu na tych samych przykładach uczących. Za każdym razem powodowało to jednak powrót sieci do przyznawania klasy większościowej dla każdego przypadku. Zdecydowaliśmy się zatem pozostać przy podejściu opisanym w poprzednim punkcie.

5. Problemy

Z dostępnym do przetwarzania zbiorem danych wiązało się niestety sporo utrudniających uczenie problemów. Po pierwsze, wiele etykiet, które miały opisywać układ dróg było w mniejszym lub większym stopniu niepoprawnych. Niektóre drogi wskazywane przez etykiety nie istniały w rzeczywistości i odwrotnie. Ponadto duża część zdjęć nie zajmowała całego rozmiaru, reszta pliku była wypełniona białym tłem. Etykiety były przy tym uzupełnione dla całego obrazu. Przykład zamieszczono na rysunku 2. Zdjęcia uznane za rażąco niepoprawne wykluczono z uczenia.



rys. 2. Przykład problematycznego zdjęcia.