State of art

Autor: Monika Etrych

W ostatnich latach kiedy mamy do czynienia z szybko rozwijającą się technologią, coraz większą mocą obliczeniową, powstają coraz szybsze sposoby znajdowania rozwiązań dla złożonych problemów. Coraz częściej w systemach i aplikacjach wykorzystywane jest rozpoznawanie różnych elementów, takich jak cyfry ręcznie pisane. Klasyfikacja cyfr znajduje zastosowanie między innymi w odczytywaniu kodów pocztowych podczas sortowania listów i przesyłek. Może być również wykorzystywana do przetwarzania danych z papierowych formularzy w bankach i urzędach [1].

Wraz z rosnącą liczbą danych wymagających systematyzowania, rośnie potrzeba dokładniejszych metod klasyfikacji. Ostatnie lata przyniosły szybki rozwój metod rozpoznawania, a w dziedzinie cyfr ręcznie pisanych osiągnięto przełomowe wyniki w zakresie dokładności. Jednak potrzeba coraz dokładniejszego klasyfikowania nadal jest bardzo ważna i w najbliższych latach ta dziedzina będzie się jeszcze znacznie rozwijać [4].

Problem porównania skuteczności wybranych metod stosowanych do rozpoznawania cyfr ręcznie pisanych podjął zespół S.M. Shamina. Przeprowadzano badania na metodach: perceptron wielowarstwowy (MLP), maszyna wektorów nośnych (SVM), naiwny klasyfikator bayesowski, sieć bayesowska, las losowy, J48 oraz drzewo decyzyjne. Spośród wybranych metod najwyższy poziom dokładności osiągnął algorytm perceptron wielowarstwowy z dokładnością na poziomie 90,37%. Najgorsze wyniki osiągnęła metoda drzewa decyzyjnego z dokładnością 85,6% [1].

W przeprowadzonych badaniach został użyty zestaw danych NIST SD19 jest to zestaw danych na podstawie którego został stworzony zestaw danych MNIST. Zestawy danych trenujących i testowych w NIST pochodziły z dwóch niezależnych źródeł, wpływa to na niedostateczne wyuczenie modelu. Kiedy model w danych uczących otrzymuje cyfry łatwe do rozpoznania, nie ma możliwości nauczenia się rozpoznawania trudniejszych przypadków, z którymi ma później styczność w zestawie testującym. W zestawie danych trenujących wykorzystuje się obrazy liczb napisane przez urzędników United States Census Bureau*,*jest to rządowa agencja, która jest odpowiedzialna między innymi za spisy ludności w Stanach Zjednoczonych. Natomiast w zestawie danych testujących wykorzystano obrazy cyfr napisane przez uczniów. W związku z problemami wynikającymi z NISTa powstał MNIST, w którym pomieszano dane z obu zestawów NISTa [1][2].

Pokazuje to dlaczego w przeprowadzonych badaniach zostały osiągnięte tak niskie wyniki, dokładność około 90%. W ostatnich latach wyuczone algorytmy na podstawie zestawu danych MNIST osiągają dokładność około 99% [3][4].

Taki wynik osiągnął między innymi Sawita Ahlawat wraz z współpracownikami, którzy podjęli próbę dopasowania jak najlepszych hiperparametrów do modelu konwolucyjnych sieci neuronowych. Sprawdzono wpływ liczby warstw, rozmiar kroku, pole widzenia sieci neuronowej, rozmiar filtra, dopełnienie(ang. padding), rozcieńczenie (ang. dilution) oraz rodzaje klasyfikatorów. Zastosowano kilka metod przygotowania danych takie jak: skalowanie, centrowanie, redukcje szumów, szacowanie przechyłu. Najlepszy wynik jaki udało się osiągnąć to dokładność na poziomie 99,89% używając optymalizatora Adam [4].

# Przypisy

[1] S. M. Shamim, Mohammad Badrul Alam Miah, Angona Sarker, Masud Rana, Abdullah Al Jobair, 2018, *Handwritten Digit Recognition Using Machine Learning,* Algorithms Indonesian Journal of Science & Technology (1) vol. 3, str.: 29-39 ISSN 2528-1410, ISSN 2527-8045

[2] Oficjalna strona MNIST dostęp: 17.01.2023, url: <http://yann.lecun.com/exdb/mnist/>

[3] A. Baldominos, Y. Saez, P. Isasi, 2019, *A Survey of Handwritten Character Recognition with MNIST and EMNIST*

[4] *S.* Ahlawat, A. Choudhary, A. Nayyar, S. Singh, B. Yoon*,* 2020, *Improved Handwritten Digit Recognition Using Convolutional Neural Networks (CNN)*