

Jedną z podstawowych właściwości obrazów cyfrowych jest ich rozdzielczość przestrzenna, wyrażana najczęściej poprzez liczbę pikseli w wymiarach poziomym i pionowym. Warto przy tym odnotować, że sama liczba pikseli w obrazie nie przesądza jeszcze, czy odpowiednio małe detale będą w obrazie możliwe do dojrzenia, bo istotną rolę odgrywają jeszcze inne czynniki związane z jakością informacji wizyjnej. Z pewnością obiektów o wielkości mniejszej niż piksel nie będziemy w stanie w obrazie rozpoznać, jednak nie jest powiedziane, że będzie to możliwe w przypadku obiektów obejmujących większy obszar. Z tego względu zwiększanie gęstości pikseli nie jest wystarczające do zapewnienia lepszej rozdzielczości rozumianej jako zdolność rozróżniania odpowiednio małych szczegółów.

Obecnie prawie każdy spotkał się z możliwością komputerowego „powiększania” małych zdjęć i obrazków, co zazwyczaj prowadzi do tego, że obraz owszem zajmuje większy obszar (czyli zawiera więcej pikseli), jednak wcale dużo więcej w nim nie widać – obraz jest zamglony, zamazany, pojawia się coś, co w języku potocznym określa się mianem „pikselozy”. Operacja taka jest wynikiem działania interpolacji, która oszacowuje kolor nowych pikseli na podstawie „starych” pikseli pochodzących z obrazu sprzed powiększania. Można sobie wyobrazić, że człowiek podszedłby do takiego zadania w bardziej twórczy sposób. Przykładowo, jeżeli w małym obrazie rozpozna jakiś obiekt, którego wygląd jest mu dobrze znany (np. znak drogowy ograniczenia prędkości), odtworzyłby jego wygląd dokładnie. Skorzystałby w tym zakresie ze swojej wiedzy o tym, że znak jest okrągły, posiada czerwoną obwódkę, a w środku znajduje się liczba. Ponadto korzystając ze swoich umiejętności plastycznych lub graficznych, byłby w stanie taki znak w obrazie powiększonym wiernie odtworzyć. Można sobie wyobrazić, że liczba widoczna w wejściowym obrazie byłaby za mała do rozpoznania – wtedy człowiek mógłby próbować ją odgadnąć, tak żeby w powiększonym obrazie pojawiła się wartość możliwa (przykładowo nie wpisałby tam liczby 58, bo wie, że na znaku prędzej pojawi się liczba 50). Ryzyko popełnienia błędu byłoby mniejsze, gdyby człowiek dysponował nie jednym małym obrazem, a wieloma, które różniąc się między sobą, przynosiłyby inną część informacji wysokorozdzielczej.

Algorytmy tak zwanej *rekonstrukcji nadrozdzielczej*, zwłaszcza te oparte o uczenie maszynowe, naśladują opisane wyżej podejście – powiększają pojedynczy obraz wejściowy lub ich serię korzystając z wiedzy o tym jak poszczególne obiekty wyglądają w niskiej i wysokiej rozdzielczości. Wiedzę tę zdobywają na podstawie odpowiednio skonstruowanego zbioru treningowego, zawierającego przedstawiające to samo sparowane obrazy o różnej rozdzielczości. W ostatnich latach obserwowaliśmy intensywny rozwój metod rekonstrukcji nadrozdzielczej, który zawdzięczaliśmy zastosowaniu spłotowych sieci neuronowych w powiązaniu z technikami uczenia głębokiego. Możliwości tych metod zdecydowanie się zwiększyły, jednak wciąż nie są one wystarczające do powszechnych zastosowań. Owszem, w wielu przypadkach pozwalają one na osiągnięcie niesamowitych efektów, zwłaszcza gdy są stosowane do obrazów, które zostały wcześniej sztucznie pomniejszone, jednak ich zachowanie nie jest wystarczająco stabilne i niezawodne, żeby mogły zostać wykorzystane w środowisku niekontrolowanym. Dotyczy to zwłaszcza przypadków, gdy na podstawie informacji zrekonstruowanej byłyby podejmowane decyzje o krytycznym znaczeniu, na przykład w obrazowaniu medycznym lub teledetekcji. W szczególności związane jest to z sytuacjami nietypowymi, gdy przykładowo w obrazie medycznym zamiast zdrowej tkanki (co ma miejsce w zdecydowanie przeważającej większości przypadków) pojawiłaby się tkanka zmieniona chorobowo – nie chcielibyśmy, żeby w obrazie zrekonstruowanym tkanka taka wyglądała na zdrową, bo mogłoby mieć to tragiczne skutki dla pacjenta. Posługując się wcześniejszym przykładem, czasami musimy być gotowi na to, że na znaku drogowym pojawi się liczba 58, mimo że takiego znaku drogowego w rejestrze oficjalnie nie ma.

Celem tego projektu jest opracowanie nowych metod rekonstrukcji o zwiększonej wiarygodności, które będą mogły być zastosowane w środowisku niekontrolowanym dla obrazów rzeczywistych (czyli nie tylko takich, które zostały sztucznie pomniejszone). Zostanie to osiągnięte poprzez zastosowanie nowej strategii treningu sieci realizujących rekonstrukcję, która będzie polegać na dążeniu do tego, żeby wynik rekonstrukcji nadawał się do realizacji różnorodnych zadań komputerowej wizji. Wśród tych zadań znajdzie się między innymi rozpoznawanie tekstu, wykrywanie małych obiektów, czy zaznaczanie zadanych obszarów (np. lasów w obrazach satelitarnych lub ludzkiej skóry w obrazach barwnych). Ponadto będziemy dążyć do opracowania technik uczenia samonadzorowanego, pozwalających na trening sieci za pomocą obrazów o niskiej i wysokiej rozdzielczości, które nie przedstawiają dokładnie tego samego. Wymaganie, żeby obrazy były sparowane (czyli przedstawiały tę samą scenę), jest bardzo trudne do zapewnienia i utrudnia gromadzenie odpowiednich ilości danych rzeczywistych. W związku z tym zniesienie lub złagodzenie tego ograniczenia ułatwi opracowanie metod rekonstrukcji nadrozdzielczej nadających się do rzeczywistych zastosowań.