

Streszczenie rozprawy doktorskiej

**Przetwarzanie i integracja wielomodalnych danych obrazowych
wspomagające wykrywanie zachowań związanych z obniżeniem poziomu
koncentracji użytkownika pojazdu mechanicznego**

Autor: mgr inż. Anton Smoliński

Promotor: dr hab. inż. Paweł Forczmański, prof. ZUT.

Promotor pomocniczy: dr inż. Adam Nowosielski

Celem niniejszej pracy doktorskiej było opracowanie metod ekstrakcji i integracji cech wielomodalnych obrazów twarzy w celu wykrywania przestąnek świadczących o obniżeniu poziomu koncentracji użytkownika pojazdu mechanicznego.

Postawiona **teza** zakłada, że odpowiednia ekstrakcja, przetwarzanie oraz integracja cech obrazów wielomodalnych twarzy pozwalają na skuteczniejszą detekcję zachowań charakterystycznych dla obniżenia poziomu koncentracji użytkownika pojazdu mechanicznego w porównaniu do metod bazujących na obrazach unimodalnych.

W ramach pracy przedstawiono autorskie metody ekstrakcji, przetwarzania i integracji obrazowych danych wielomodalnych mające na celu wykrywanie przestąnek świadczących o obniżeniu poziomu koncentracji użytkownika pojazdu mechanicznego. Przeprowadzono badania eksperymentalne i dokonano oceny różnych wariantów ekstrakcji oraz fuzji danych na różnych etapach przetwarzania. Wykorzystano dane z różnych modalności obrazowych, takich jak obrazy RGB, obrazy termowizyjne oraz mapy głębokości, co stanowi rozszerzenie w stosunku do tradycyjnych rozwiązań opartych wyłącznie na pojedynczych modalnościach. Przeanalizowano zarówno tradycyjne metody ekstrakcji cech i rozpoznawania wzorców, jak i techniki oparte na głębokim uczeniu maszynowym, w tym sztuczne sieci neuronowe CNN oraz LSTM, umożliwiające analizę dynamicznych zmian w zachowaniu kierowcy.

Szczególną uwagę poświęcono detekcji oznak zmęczenia i rozproszenia uwagi, takich jak częstotliwość mrugania, ziewanie, opadanie głowy czy pocieranie oczu. Wprowadzono wskaźnik PEROPEN, służący do pomiaru poziomu zmęczenia na podstawie analizy stopnia otwarcia ust. Ponadto, zbadano wpływ metod detekcji twarzy, oczu i ust na efektywność oceny poziomu koncentracji, co przyczyniło się do optymalizacji procesu detekcji.

Wyniki badań wykazały, że integracja wielomodalnych danych obrazowych znacząco poprawia skuteczność detekcji wizualnych przestąnek świadczących o zmęczeniu oraz rozproszeniu uwagi kierowcy w porównaniu z analizą pojedynczych modalności, co potwierdza postawioną tezę.

W ramach prowadzonych badań stworzono i udostępniono zbiór testowy w serwisie cvlab.zut.edu.pl, zawierający zarówno surowe, jak i przetworzone dane do badań. Umożliwia to innym badaczom wykorzystanie tych zasobów do testowania własnych metod i algorytmów.

Wnioski z przeprowadzonych badań wskazują, że opracowane metody mogą przyczynić się do stworzenia zaawansowanych i precyzyjnych systemów monitorowania stanu kierowców. Mogą one znaleźć zastosowanie zarówno w nowoczesnych, jak i starszych pojazdach, zwiększając tym samym bezpieczeństwo na drogach.

Słowa kluczowe: widzenie komputerowe, przetwarzanie obrazów, obrazowanie wielomodalne, wykrywanie zmęczenia kierowców, uczenie maszynowe, fuzja danych.

17.09.24

Anton

Smoliński