

Wizyjny system rozpoznawania i identyfikacji cen paliw

KATEDRA PRZETWARZANIA SYGNAŁÓW I INŻYNIERII MULTIMEDIALNEJ

autor: mgr inż. Jędrzej Łukasiuk, teleinformatyka S2, rok ukończenia 2018

opiekun pracy: dr hab. inż. Krzysztof Okarma, prof. nadzw. ZUT

Wstęp

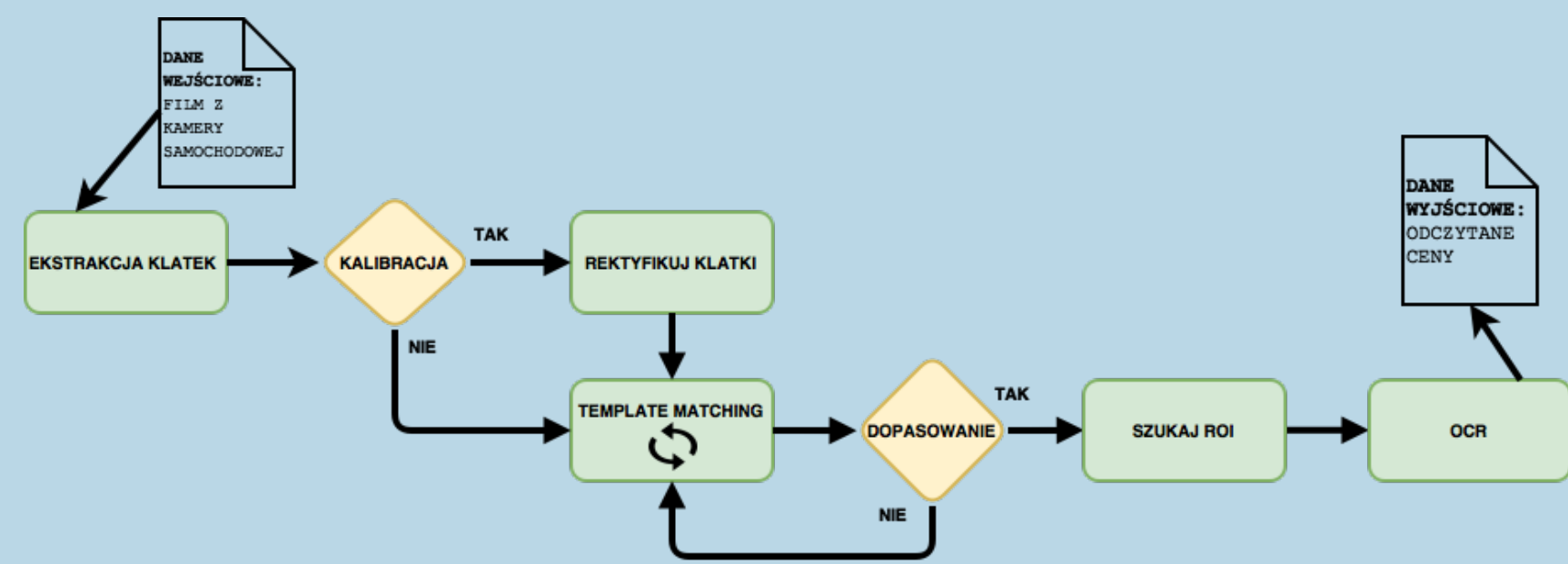
Celem pracy było opracowanie systemu wizyjnego służącego do rozpoznawania cen paliw umieszczonych na banerach znajdujących się przy stacjach benzynowych, na podstawie filmów pozyskiwanych z kamery umieszczonej wewnątrz samochodu.

Opis problemu

Opracowanie systemu działającego w niedeterministycznych warunkach wymaga integracji kilku niezależnych technik, do których możemy zaliczyć:

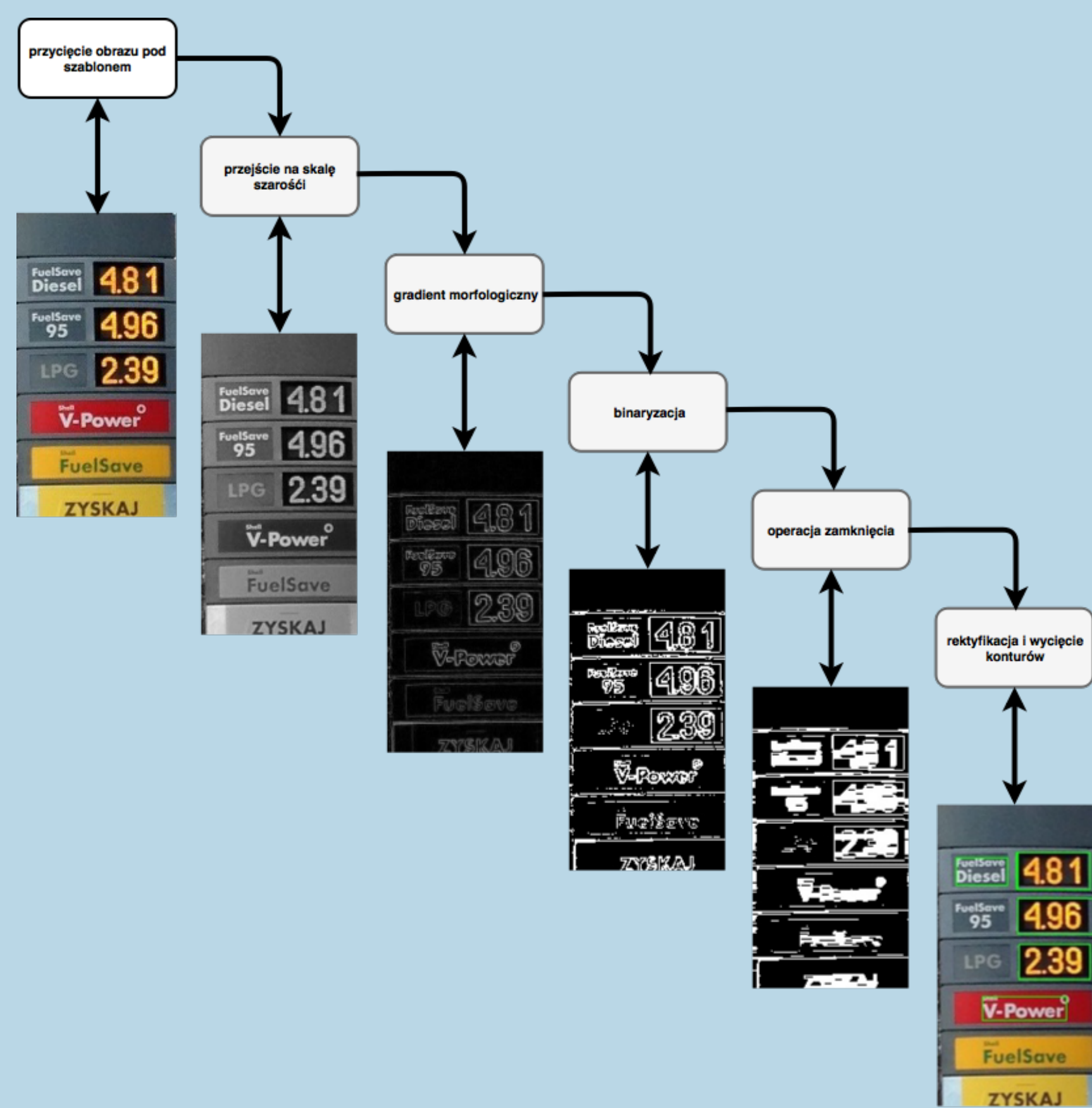
- **template matching** – technika dopasowywania obrazu do zadanego wzorca,
- **przekształcenia morfologiczne** – grupa operacji służąca przede wszystkim do analizy kształtów na obrazach binarnych,
- **kalibracja kamery** – proces pozyskiwania parametrów wewnętrznych kamery i parametrów dystrybucji. Dzięki niemu możemy rektyfikować obraz (poprawiać jego parametry geometryczne),
- **OCR (ang. Optical Character Recognition)** – zestaw technik służących do rozpoznawania znaków na obrazach rastrowych.

Uproszczony algorytm działania systemu wygląda następująco:



Wyszukiwanie ROI

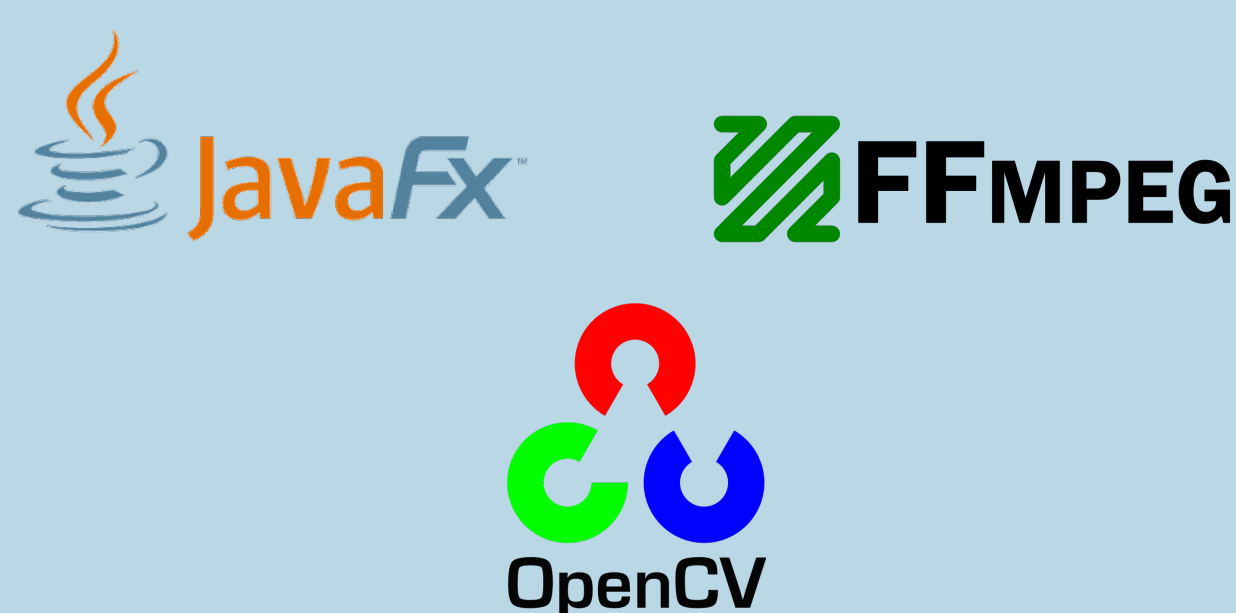
Kluczowym etapem algorytmu jest wyszukiwanie obszarów zainteresowania (ang. ROI – Region of Interest). W moim przypadku będą to te fragmenty obrazu, na których znajdują się ceny paliw. W tym celu przekształcono klatkę filmu, na którym znaleziono logotyp stacji zgodnie z zależnościami podanymi na rysunku poniżej.



Zakreślone na powyższym obrazie zielone ramki są wycinane, a następnie podawane do silnika OCR.

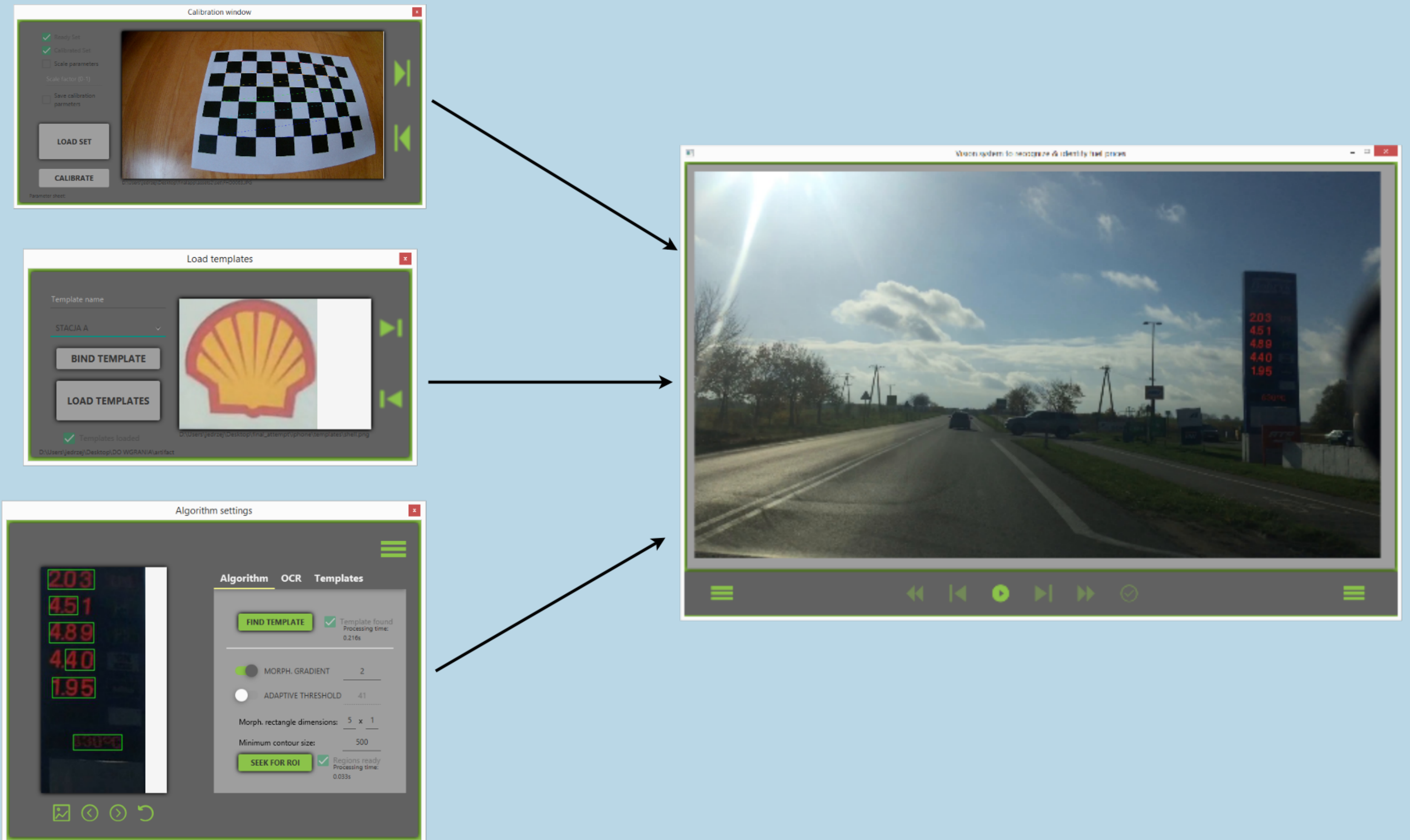
Stos technologiczny

Aplikację użytkową napisano w języku *Java* (z wykorzystaniem najnowszej biblioteki graficznej *JavaFX*). Do przetwarzania obrazów skorzystano z biblioteki *OpenCV*, a możliwość manipulacji filmami zapewnił program *FFmpeg*.



Aplikacja użytkowa

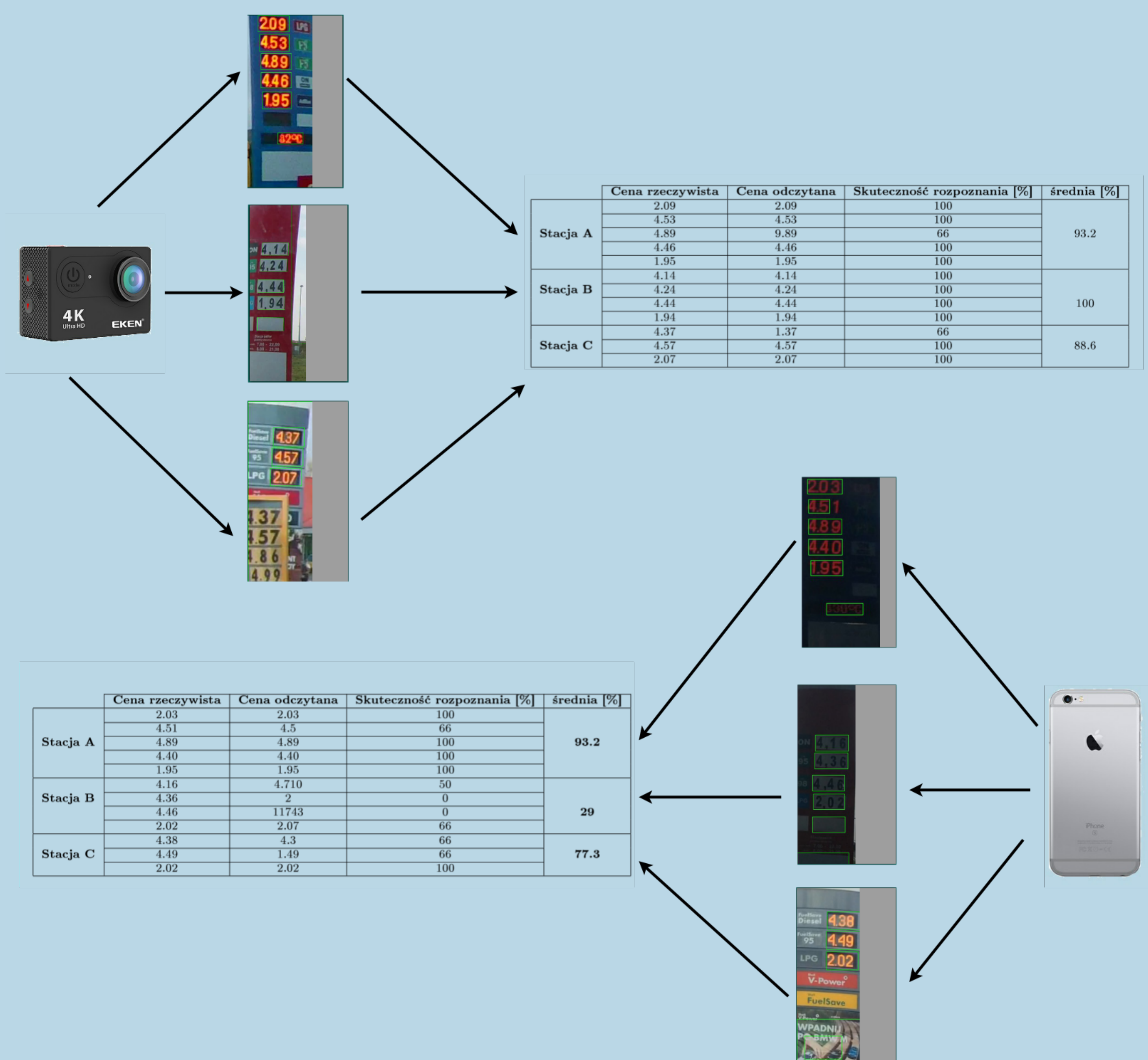
Aplikacja składa się z głównego panelu, wraz z 3 oknami modalnymi służącymi do kalibracji kamery, ładowania szablonów oraz trenowania algorytmu.



Użytkownik może przeszukiwać film w trybie automatycznym i ręcznym. W pierwszym przypadku należy załadować szablony w postaci emblematów stacji i nacisnąć przycisk start. Program automatycznie sprawdzi każdą klatkę wykorzystując do tego architekturę wielowątkową (*CountDownLatch* – jeden wątek na jeden logotyp). W drugim przypadku użytkownik ręcznie przeszukuje film przesuując się po poszczególnych klatkach.

Testowanie algorytmu. Wnioski

Algorytm przetestowano dla 2 kamer, z których łącznie pozyskano 6 filmów (po 3 na każdą z nich, przedstawiające te same stacje). Podczas testowania systemu algorytm trenowano ręcznie celem uzyskania jak najlepszych rezultatów.



WNIOSKI

W przypadku kamery Eken uzyskano średnią skuteczność rozpoznania na poziomie 94 %, natomiast dla kamery iSight wynosiła ona 66.5 %. Skuteczność identyfikacji regionów zainteresowania wyniosła łącznie 100 %. Jedynym przypadkiem, w którym rozpoznanie cen nie przyniosło oczekiwanych rezultatów było zdjęcie z niewystarczającym oświetleniem (kolor tła zlewał się z ceną). Poza tym system mylił się dla pojedynczych cyfr, których struktura jest podobna (np. "1" i "4"), co jest specyfiką samego silnika OCR.