



Fundusze Europejskie
dla Mazowsza



Rzeczpospolita
Polska

Dofinansowane przez
Unię Europejską



Mazowsze.
serce Polski



ARTEX IT

Końcowy Raport z Realizacji Zadania Badawczego

Nazwa podmiotu realizującego projekt: "ARTEX IT" ARTUR GREGORCZYK

Numer projektu: FEMA.01.01-IP.01-02IN/24

Tytuł projektu: System identyfikacji zachowań pracowników za pomocą wizji z wykorzystaniem modeli AI

Okres sprawozdawczy: 01.09.2025 – 28.02.2026

1. Wprowadzenie

Celem niniejszego raportu końcowego jest przedstawienie wyników prac badawczych zrealizowanych w ramach Zadania III, pt.: „Opracowanie w ramach prac rozwojowych metod AI do analizy i fuzji danych z poszczególnych kamer wizyjnych oraz innych danych pomocniczych za pomocą sieci neuronowej”. Prace stanowią bezpośrednią kontynuację działań przeprowadzonych w poprzednich etapach projektu.

Hipoteza badawcza Zadania III zakłada, że integracja i analiza danych z wielu kamer wizyjnych oraz pomocniczych źródeł danych (ERP, RFID/iBeacon) za pomocą algorytmów sztucznej inteligencji — w szczególności sieci neuronowych — znacząco zwiększy dokładność pomiaru liczby osób i ich rozmieszczenia w załączonych przestrzeniach produkcyjnych.

Kluczowymi wskaźnikami weryfikacji są:

Wskaźnik 1 Skuteczność pomiaru czasu przebywania w strefie na poziomie min. 90%	Wskaźnik 2 Skuteczność ogólnej liczby osób ze wszystkich kamer na poziomie min. 95%
---	---

2. Opis Zaplanowanych Prac Badawczych

W okresie przewidzianym harmonogramem realizację Zadania III podzielono na pięć etapów badawczych zgodnych ze specyfikacją zadania:

01	02	03
Etap 1 Rozwój algorytmów AI do przetwarzania danych wizyjnych i z innych systemów	Etap 2 Integracja i synchronizacja danych z kamer, systemów ERP i RFID/iBeacon	Etap 3 Analiza i interpretacja danych — identyfikacja wzorów zachowań, ocena ryzyka, predykcja scenariuszy
04	05	
Etap 4 Testowanie i walidacja systemu w warunkach kontrolowanych i rzeczywistych	Etap 5 Optymalizacja systemu oraz testowe wdrożenie z monitorowaniem na hali produkcyjnej	

3. Zakres Zrealizowanych Działań

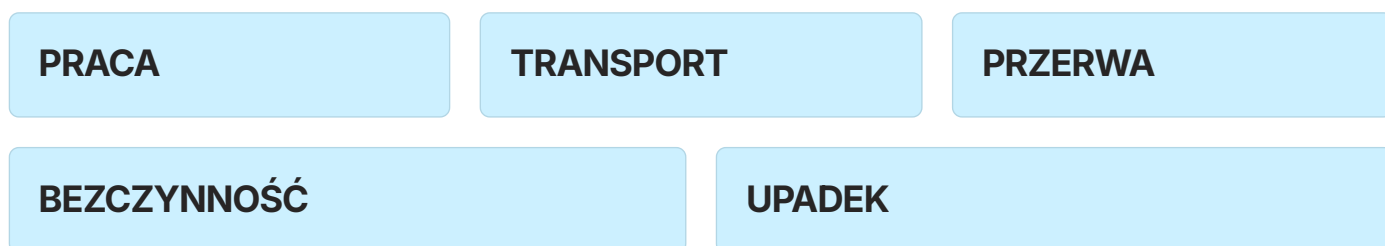
3.1 Architektura Systemu – Moduł Wizyjny (Warstwa Detekcji i Śledzenia)

System zaimplementowano w języku **Python** z wykorzystaniem bibliotek **OpenCV** i **bibliotek do detekcji obiektów**. Rdzeń detekcji oparty jest na **modelu detekcji pozy** z estymacją **17-punktowego szkieletu ciała** (standard COCO). Próg pewności detekcji ustalono na poziomie **CONF = 0,65**, co istotnie redukuje liczbę fałszywych alarmów w zatłoczonym środowisku produkcyjnym.

Do ciągłego śledzenia osób zastosowano **algorytm śledzenia wieloobjektowego** uzupełniony o **filtr Kalmana** (wygładzanie trajektorii i estymacja prędkości). Przekształcenie perspektywiczne (homografia) umożliwia wyrażanie pozycji pracowników w układzie metrycznym **[m]**, stanowiąc podstawę obliczeń gęstości i strefowania.

3.2 Maszyna Stanów (FSM) i Analiza Biomechaniczna

Każdy śledzony pracownik posiada własną instancję automatu skończonego (FSM) klasyfikującego jego aktywność do jednej z pięciu klas:



Klasyfikacja opiera się na trzech sygnałach biomechanicznych: prędkości translacyjnej (filtr Kalmana), kącie stawu kolanowego oraz kącie pochylenia tułowia. Zmiany stanów buforowane są przez filtrację czasową (15 klatek, próg 70% zgodności), eliminując artefakty chwilowych błędów detekcji. Wyjątkiem jest stan **UPADEK** — ze względów bezpieczeństwa powoduje natychmiastową zmianę stanu bez opóźnienia.

3.3 Wielokryterialna Detekcja Upadków

W systemie zaimplementowano wielokryterialny algorytm wykrywania upadków oparty na trzech niezależnych sygnałach, które muszą wystąpić łącznie — co najmniej dwa z trzech — by uruchomić detekcję kandydatury:

Kryterium A – Proporcja ramki detekcji $h/w < 0,60$ — leżąca osoba ma ramkę szerszą niż wyższą	Kryterium B – Kąt pochylenia tułowia > 55° względem pionu (pozycja pozioma)	Kryterium C – Nagły spadek pozycji bioder Znormalizowany spadek > 15% wysokości ramki detekcji między kolejnymi klatkami
--	---	--

Potwierdzenie upadku wymaga utrzymania stanu kandydatury przez zdefiniowane okno czasowe, co skutecznie eliminuje fałszywe alarmy (ang. *false positives*) wywołane standardowymi czynnościami ruchowymi, takimi jak kucanie lub schylenie. Wyjście z detekcji upadku (odwołanie alarmu) następuje automatycznie w momencie, gdy geometryczna weryfikacja sylwetki — oparta na analizie proporcji ramki otaczającej oraz znormalizowanym kącie pochylenia tułowia — wskaże na ustabilizowany powrót pracownika do pozycji stojącej.

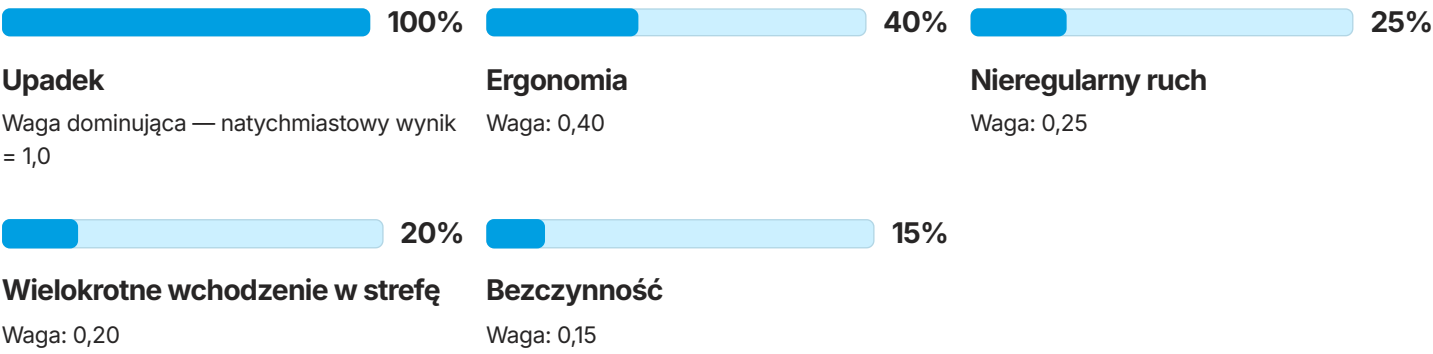
☒ W interfejsie wizualnym i module analitycznym incydent ten traktowany jest priorytetowo: sygnalizowany jest **migającą czerwoną ramką**, komunikatem ostrzegawczym „**UPADEK**” oraz przypisaniem maksymalnego współczynnika ryzyka (`risk_score = 1,0`).

3.4 Warstwa Analityczna – Analiza Behawioralna i Ocena Ryzyka

System wyposażono w wielopoziomową warstwę analityczną uruchamianą cyklicznie co **1 sekundę** (`ANALYSIS_INTERVAL = 1,0 s`), co istotnie przewyższa wymagany wskaźnik aktualizacji (co 1 minutę). Warstwę tworzą cztery współpracujące moduły:

BehaviorAnalyzer Analiza per-pracownik — pięć wzorów behawioralnych: - Przedłużona bezczynność (<code>IDLE > 30 s</code>) — WARNING/ALERT - Przedłużona przerwa (<code>BREAK > 120 s</code>) - Ryzyko ergonomiczne — pochylenie > 30° przez > 5 s - Niestabilny wzorec ruchu — odchylenie std. prędkości > 0,40 m/s - Wielokrotne wchodzenie w strefę — > 4 wejść w oknie 120 s	SceneAnalyzer Analiza sceny globalnej: - Gęstość zaludnienia strefy (os/m^2) z progiem alertu 0,08 os/m^2 - Izolacja pracownika — jedna osoba w strefie przez > 60 s - Zsynchronizowane przerwy — > 50% pracowników jednocześnie w stanie PRZERWA - Trend gęstości (rosnący / malejący / stabilny) na podstawie okien 10-sekundowych
	ScenarioPredictor Predykcja krótkoterminowa (horyzont 30 s) na podstawie ekstrapolacji liniowej okna analizy trendu (<code>TREND_WINDOW = 60 s</code>); identyfikuje potencjalną eskalację zagęszczenia, zmęczenie załogi (udział <code>IDLE > 40%</code>) oraz systemowe ryzyko ergonomiczne (> 30% aktywnych z pogorszoną postawą).

Wieloczynnikowy współczynnik ryzyka (`risk_score`) w przedziale [0,0; 1,0] per-pracownik obliczany jest jako ważona suma pięciu czynników:



W interfejsie wizualnym poziom ryzyka $r > 0,30$ sygnalizowany jest czerwoną ramką grubości proporcjonalnej do wartości wskaźnika.

3.4 Optymalizacja i Adjustowanie Modelu

Po uzyskaniu wstępnych wyników testowych przystąpiono do procesu optymalizacji modelu w celu poprawy jego stabilności i zdolności generalizacyjnych. Kluczowym elementem tego etapu było zastosowanie walidacji krzyżowej, która pozwoliła ocenić rzeczywistą skuteczność modelu w różnych warunkach operacyjnych oraz zidentyfikować potencjalne problemy z przeuczeniem. Zbiór danych podzielono na pięć niezależnych podzbiorów i przeprowadzono pełny cykl treningu oraz walidacji dla każdej z kombinacji, obliczając średnie wartości metryk skuteczności.



Podzbiory walidacji krzyżowej

Cross-validation

Pełny cykl treningu i walidacji dla każdej kombinacji

Obliczanie średnich wartości metryk skuteczności

Równolegle rozszerzono strategię augmentacji danych, wprowadzając dodatkowe przekształcenia symulujące zmienne warunki oświetleniowe charakterystyczne dla środowiska przemysłowego, w tym symulację cieni, refleksów oraz różnych pór dnia. Zaimplementowano również augmentacje geometryczne uwzględniające różne kąty widzenia oraz zakłócenia tła typowe dla hal produkcyjnych.

Walidacja krzyżowa (5-fold)
Ocena stabilności modelu w różnych warunkach operacyjnych

Augmentacja danych
Symulacja cieni, refleksów, różnych pór dnia, kątów widzenia

Regularyzacja
Techniki ograniczające ryzyko nadmiernego dopasowania do zbioru treningowego

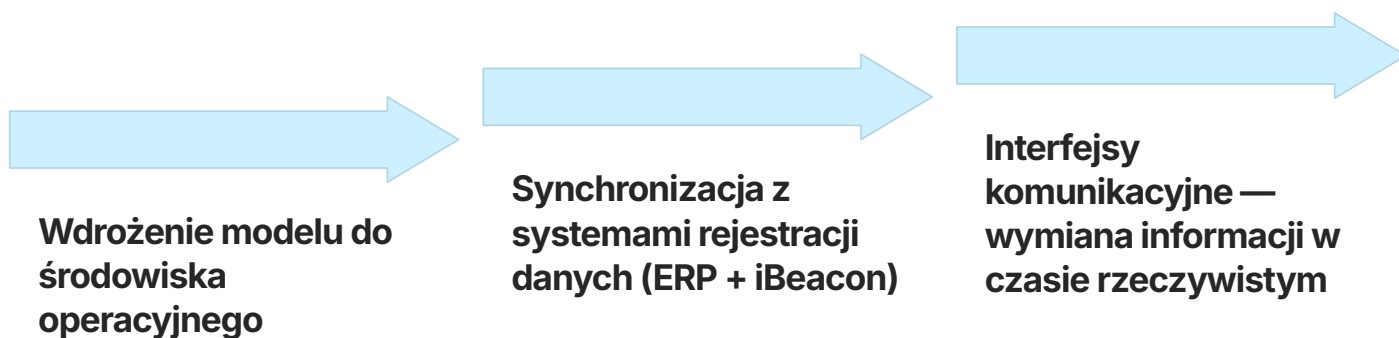
Dostrojenie hiperparametrów
Learning rate, rozmiar batcha, liczba epok

Wyniki walidacji krzyżowej potwierdziły stabilność modelu oraz jego zdolność do skutecznej pracy w zróżnicowanych warunkach operacyjnych występujących w rzeczywistym środowisku produkcyjnym.

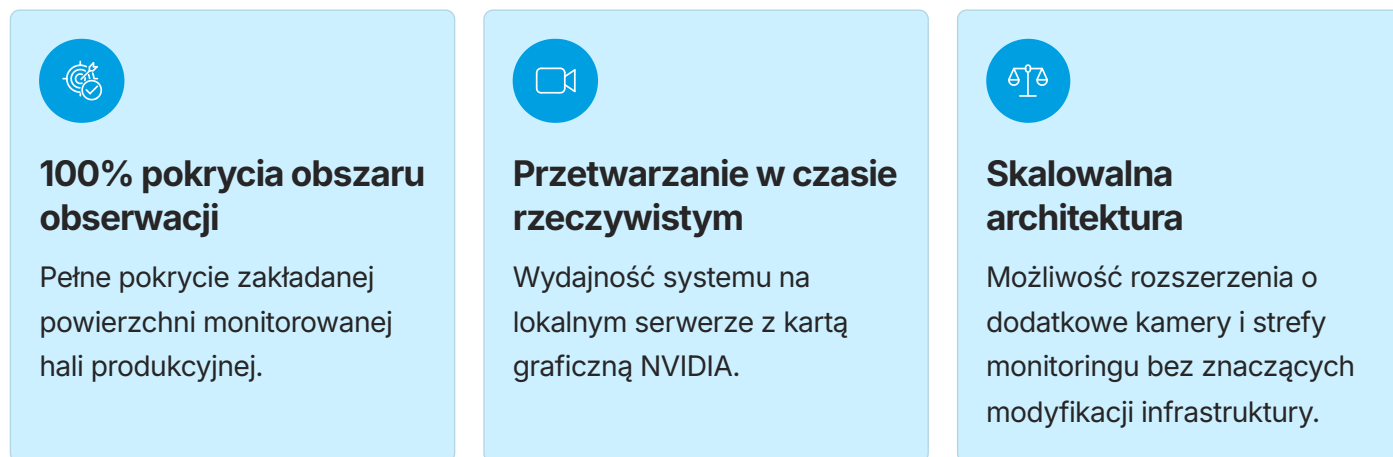
Stabilność modelu	Zdolność generalizacyjna
Potwierdzona przez walidację krzyżową w zróżnicowanych warunkach operacyjnych	Skuteczna praca w rzeczywistym środowisku produkcyjnym

3.5 Skalowanie i Integracja Modelu – Integracja z Systemem Monitoringu

Ostateczny system został zintegrowany z istniejącą infrastrukturą monitoringu hali produkcyjnej, tworząc kompletne rozwiązanie do analizy danych w czasie rzeczywistym. Proces integracji obejmował wdrożenie modelu do środowiska operacyjnego, synchronizację z systemami rejestracji danych oraz utworzenie interfejsów komunikacyjnych umożliwiających wymianę informacji z bazą danych ERP i systemem iBeacon.



Przeprowadzone testy operacyjne w warunkach rzeczywistego środowiska produkcyjnego potwierdziły prawidłowe działanie systemu oraz jego wydajność na poziomie umożliwiającym przetwarzanie strumieni wideo w czasie rzeczywistym — wymagało to wykorzystania lokalnego serwera z kartą graficzną NVIDIA. System osiągnął pełne pokrycie zakładanego obszaru obserwacji, spełniając tym samym jeden z kluczowych wskaźników sukcesu projektu.



Zaimplementowano również mechanizmy monitorowania wydajności i jakości działania systemu, które umożliwiają bieżącą kontrolę jego parametrów operacyjnych oraz wczesne wykrywanie potencjalnych problemów technicznych. Architektura rozwiązania została zaprojektowana z uwzględnieniem możliwości dalszego skalowania.

NVIDIA GPU

Serwer lokalny

Wymagany do przetwarzania strumieni wideo w czasie rzeczywistym

100%

Pokrycie obszaru

Pełne pokrycie zakładanej powierzchni obserwacji

3.5 Interfejs Użytkownika i Wizualizacja Danych

Opracowany demonstrator został wyposażony w zintegrowany, graficzny interfejs czasu rzeczywistego (interfejs nakładkowy, ang. *Heads-Up Display* – HUD), zaimplementowany z wykorzystaniem biblioteki do przetwarzania obrazu.

W warstwie wizualnej każdy śledzony pracownik reprezentowany jest przez następujące elementy:

		
Ramka otaczająca (bounding box)	17-punktowy model szkieletowy	Etykieta śledząca
Kolor dynamicznie odpowiada bieżącemu stanowi: czerwony = PRACA, pomarańczowy = TRANSPORT, żółty = PRZERWA, szary = BEZCZYNNOŚĆ, migający czerwony = UPADEK.	Uwidacznia anatomiczne punkty kluczowe (ang. <i>keypoints</i>) w kolorze zielonym.	Zawiera: unikalny identyfikator (ID), nazwę bieżącego stanu behawioralnego, prędkość translacyjną [m/s], aktualny współczynnik ryzyka, flagę obecności w strefie przestrzennej oraz skumulowany czas i liczbę wejść do strefy.

Zdefiniowana strefa zainteresowania (ROI) wizualizowana jest na płaszczyźnie hali jako wielokąt narysowany linią przerywaną z ponumerowanymi wierzchołkami.

Główny panel analityczny (górna sekcja interfejsu nakładkowego) wyświetla w czasie rzeczywistym kluczowe metryki sesji: liczbę osób przebywających w strefie, łączną liczbę aktywnych torów śledzenia, liczbę unikalnych osób rozpoznanych w kadrze, liczbę odrzuconych fałszywych detekcji w bieżącej klatce oraz informację o aktywnym trybie optymalizacji sprzętowej (przetwarzanie co N klatek). Poniżej prezentowana jest bieżąca gęstość zaludnienia strefy [os./m²] wraz z estymowanym trendem.

Dodatkowo, w lewym dolnym rogu ekranu, dynamicznie wyświetlane są trzy ostatnie zdarzenia analityczne, wykorzystujące kodowanie kolorystyczne odpowiadające poziomowi ważności incydentu:

INFO INFORMACJA	WARNING OSTRZEŻENIE
ALERT ALARM	CRITICAL KRYTYCZNY

3.6 Eksport Danych

System generuje cztery ustrukturyzowane formaty plików, zapewniając pełną identyfikowalność danych:

states_log.csv	Próbkowanie co 5 klatek — niskopoziomowe dane śledzenia: ID, stan behawioralny, prędkość, pozycja metryczna [m] oraz wskaźnik risk score.
*_vision_stream.jsonl	1 rekord/tor/s — stan pracownika uzupełniony o metryki jakości; stanowi główny strumień do późnej fuzji danych.
*_daily_summary.csv	Agregacja per sesja — skumulowane podsumowanie sesji (m.in. czas w poszczególnych stanach, wskaźniki utraty i fragmentacji toru, wystąpienie upadku).
*_events.csv	Rejestr asynchroniczny — dziennik incydentów analitycznych (znacznik czasu, kategoria zdarzenia, waga/priorytet oraz risk_score).

3.7 Integracja z Systemami ERP i RFID/iBeacon

Integracje z zewnętrznymi systemami rozwiązano na dwa sposoby:

1

REST API – Stremsoft

Wykorzystano standardowe REST API producenta oprogramowania, które pozwala pobierać zlecenia produkcyjne, meldunki produkcyjne, komplet danych z metodologii produkcji jak również całego systemu. Stworzono mechanizm webhook API łącznie z pełną dokumentacją do łączenia się zewnętrznymi systemów — rozbudowano o **202 endpointy REST API**.

2

Połączenie do bazy pośredniczącej (SQL)

System został rozbudowany o możliwość integracji na poziomie bazy danych SQL, tak że firma nie posiadająca API ani niekorzystająca z API jedynie może synchronizować dane z bazą SQL (OPC serwer).

Dodano integrację kart **RFID** — jako logowanie do panelu, aby jednoznacznie przypisać wykonanie pracy przez dany obszar. Kartę RFID wykorzystano do identyfikacji obszaru. Integracja z **iBeacon** służyła do potwierdzania przebywania w danym obszarze ilości osób. System ze względu na **RODO** nie rejestruje danych osobowych.

Synchronizacja danych następuje co **1 minutę** i ustawienie takiego czasu nie spowodowało żadnych opóźnień w bazach danych ani problemów. System webhooków jest przygotowany na wymianę danych w czasie rzeczywistym. System ERP/MES dostarcza danych odnośnie aktualnie prowadzonej produkcji — zlecenia produkcyjne i meldunki.

3.8 Testowe Wdrożenie i Monitoring

Wdrożenie na hali produkcyjnej odbyło się w **Cukierni**, gdzie przez **1 dzień produkcyjny** była weryfikacja systemu. Ocena była na **12 osobach** na hali produkcyjnej. Kamery realizowały streaming obrazu wideo na serwery GPU, gdzie następowała analiza wizyjna.

- ❑ Napotkaliśmy na problem podczas testów: w czasie testów zidentyfikowano problem osób pojawiających się na skraju obszaru roboczego, co zniwelowano systemem filtrów czasowych — oddano korektę w kodzie, który krótkie (do 5 sekund) wyjścia zupełnie pomija.

4. Weryfikacja Wskaźników Zadania

Na rys. 1–3 przedstawiono przykładowe widoki z pracy systemu w cyklu dziennym na testowym miejscu.



Rys. 1–3. Przykłady działania rozwiązania na hali testowej

4.1 Skuteczność Detekcji i Zliczania Osób

System wyposażony jest w wielowarstwowy mechanizm filtracji fałszywych alarmów, którego celem jest zbliżenie precyzji detekcji do wymaganych progów: **95%** (liczba osób) i **90%** (czas przebywania w strefie). Kluczowe składowe mechanizmu:

- Podniesiony próg pewności modelu (CONF = 0,65 vs domyślne 0,50)
- Filtr geometryczny punktów kluczowych (*keypoints*); wymagane min. 5 widocznych punktów
- Filtr proporcji ramki detekcji (*bbox*) ($h/w \geq 0,80$) z wyłączeniem dla potwierdzonych upadków
- Filtr anatomiczny: obecność przynajmniej barku lub biodra

4.2 Częstotliwość Aktualizacji Stanu

Wymaganie określa aktualizację co **1 minutę** (60 sekund). System realizuje aktualizację warstwy analitycznej (stany, anomalie, współczynnik ryzyka) co **1 sekundę** (ANALYSIS_INTERVAL = 1,0 s). Moduł strumienia wizyjnego (VisionStreamWriter) zapisuje dane per-tor z rozdzielczością 1 rekordu/sekundę. Interfejs nakładkowy (HUD) aktualizuje się synchronicznie z przetwarzanymi klatkami, średnio **25–30 razy na sekundę** przy pełnym przetwarzaniu każdej klatki. Tym samym wskaźnik jest spełniony z istotnym marginesem bezpieczeństwa.