

1) Wynikiem prac zrealizowanych w ramach Fazy I powinno być w szczególności:

1. Stworzenie zbiorów danych obrazowania metodą przemysłowej TK oraz RP, służących do trenowania modeli uczenia maszynowego z uwzględnieniem informacji o rzeczywistych wymiarach elementów w ramach badanych obiektów. Każdy zbiór będzie się składał z par: obraz TK lub RP oraz uzyskana metodą referencyjną (w przypadku komór jonizacyjnych – pomiar maszyną współrzędnościową, w przypadku puszek aerozolowych – metoda wagowa) odpowiadającą temu obrazowi wartość objętości, która w procesie uczenia będzie uznawana za wartość prawdziwą. Minimalny wolumen danych dla zbiorów treningowych i testowych:
 - a. Treningowego i testowego zbioru danych dla komór jonizacyjnych wraz z informacją o oczekiwanej wartości predykcji objętości oraz ilości elementów montażowych o odpowiadających różnej gęstości w obrazowaniu TK:
 - i. Minimalna liczebność zbioru treningowego: 250
 - ii. Minimalna liczebność zbioru testowego: 50
 - b. Treningowego i testowego zbioru danych dla pojemników aerozolowych wraz z informacją o oczekiwanej wartości predykcji objętości:
 - i. Minimalna liczebność zbioru treningowego: 250
 - ii. Minimalna liczebność zbioru testowego: 50
2. Stworzenie wstępnej wersji systemu, która może zostać poddana ocenie – opracowanie wymagań oraz architektury systemu teleinformatycznego, a także pierwszego prototypu aplikacji (nie uwzględniając integracji z modelami sztucznej inteligencji).
3. Wstępnie wytrenowany model detekcji na podstawie danych syntetycznych, które odwzorowują dane rzeczywiste.

Faza I projektu będzie poświęcona opracowaniu referencyjnego zbioru danych oraz stworzeniu wstępnej wersji systemu i modelu detekcji. Budżet realizacji fazy I zakłada pokrycie kosztów związanych z zatrudnieniem pracowników (min. naukowych i inżynierów), pracami nad wymiarowaniem i pozyskaniem obiektów będących podstawą obrazowania oraz pokryciem kosztów pośrednich.

Obiekty (komory jonizacyjne oraz zdjęcia puszek referencyjnych oraz badanych) wymagane do stworzenia niniejszych zbiorów danych zostaną dostarczone przez Główny Urząd Miar. Dane nie są objęte prawnymi ograniczeniami dostępu.

Biorąc pod uwagę dwójakie przeznaczenie rezultatów projektu zbiór ten musi dotyczyć obu rozważanych przypadków tj. komór jonizacyjnych (aplikacja radiologiczna) oraz puszek o różnym stopniu napelnienia (kontrola produktów aerozolowych) – referencyjnych oraz badanych.

Format danych obrazowych będzie zależny od rodzaju obrazów PR/TK oraz od narzędzi używanych do ich przechowywania i przetwarzania. Może to być np. format DICOM (stosowany dla obrazów medycznych) lub

Załącznik nr 13. Wytyczne dot. efektów realizacji kolejnych faz projektu.

formaty plików obrazów cyfrowych dla obrazów pochodzących z innych źródeł. Format danych dotyczących wymiarów elementów na obrazach może mieć postać tekstową, zawierając informacje o wymiarach elementów składowych obrazowanych obiektów (np. format CSV). Jeden przypadek będzie więc opisywany parą plików dot. obrazowania oraz wymiarowania.

2) Wynikiem prac zrealizowanych w ramach fazy II powinno być w szczególności:

1. Opracowanie modeli i algorytmów wykorzystujących narzędzia sztucznej inteligencji spełniających normy opisane w wymaganiach dla realizowanych projektów do precyzyjnego wyznaczania objętości elementów o różnych gęstościach w ramach badanych obiektów. Ocena opracowanych algorytmów będzie odbywać się na podstawie takich miar jak: wydajność (czas wykonania predykcji), dokładność (różnica między wartością uznaną za rzeczywistą a wartością przewidzianą przez model) oraz precyzja (zdolność modelu do powtarzalności wyników). Oczekiwane rozwiązanie w postaci rozwiązania będącego w stanie dokonać segmentacji elementów o różnej gęstości jest oparte o uczenie nienadzorowane (segmentacja elementów składowych obiektu zamkniętego na podstawie informacji o ilości elementów montażowych w zbiorze danych).
 - a. Wydajność – czas wykonania predykcji dla danego obiektu nie powinien przekraczać 30 sekund.
 - b. Dokładność – wymagana dokładność w wyznaczaniu objętości:
 - i. poniżej 0,4% w przypadku wyrobów aerozolowych;
 - ii. poniżej 0,1% w przypadku detektorów promieniowania.
 - c. Precyzja – powtarzalność uzyskanych wyników dla tych samych przedmiotów poddanych analizie mieści się poniżej 0,1%. Na podstawie mapy segmentacji elementów, rozwiązanie jest w stanie wyznaczyć i zaprezentować ich wymiary (a także objętość).
2. Opracowanie prototypu rozwiązania w formie narzędzia informatycznego (aplikacji webowej), spełniającego wymagania funkcjonalne względem przypadków użycia, tj. komór jonizacyjnych (aplikacja radiologiczna) oraz puszek o różnym stopniu napełnienia (kontrola produktów aerozolowych), i umożliwiają integrację z modelami sztucznej inteligencji, a także umożliwiającego pozyskiwanie informacji o charakterystyce segmentowanych elementów o różnych gęstościach z uwzględnieniem ich objętości na podstawie obrazowania obiektów metodą przemysłowej tomografii komputerowej. Minimalne oczekiwane funkcjonalności aplikacji umożliwiają uzyskanie predykcji z systemu na podstawie przesłanego obrazowania za pośrednictwem formularza na platformie.

Faza II projektu będzie poświęcona opracowaniu modeli i algorytmów uczenia maszynowego wykorzystując dane pozyskane w fazie I projektu oraz przygotowaniu prototypu docelowego rozwiązania. Opracowane rozwiązania zostaną porównane między sobą pod względem wydajności, dokładności i precyzji, a także skuteczności w wyznaczaniu objętości elementów o różnych gęstościach. Budżet realizacji fazy II zakłada pokrycie kosztów

Załącznik nr 13. Wytyczne dot. efektów realizacji kolejnych faz projektu.

związanych z zatrudnieniem pracowników (min. naukowych i inżynierów), wydzieleniem potrzebnych zasobów obliczeniowych oraz pokryciem kosztów pośrednich.

3) Wynikiem prac zrealizowanych w ramach fazy III powinno być w szczególności:

1. Opublikowanie wyników prac uzyskanych w fazie II nad zastosowaniem metod sztucznej inteligencji w ramach rozwiązania będącego wynikiem prac projektu.
2. Opracowanie finalnej wersji rozwiązania w formie narzędzia informatycznego (aplikacji webowej) umożliwiającego pozyskiwanie informacji o charakterystyce segmentowanych elementów o różnych gęstościach z uwzględnieniem ich objętości na podstawie obrazowania obiektów metodą przemysłowej tomografii komputerowej. Główne funkcjonalności systemu umożliwiają przeprowadzania analiz obrazowania CT nowych obiektów z wykorzystaniem powyższych modeli.
 - a. Narzędzie powinno udostępniać funkcjonalności za pośrednictwem graficznego interfejsu użytkownika (GUI) a także poprzez dedykowane usługi sieciowe umożliwiające integrację zewnętrznych narzędzi w celu uzyskania predykcji.
 - b. Narzędzie powinno integrować mechanizmy kontroli dostępu na podstawie upoważnionych kont oraz autoryzowanego dostępu do interfejsu programistycznego aplikacji.
3. Dokumentacja modeli uczenia maszynowego wraz z informacją dot. użytych algorytmów, zdefiniowanej funkcji kosztu, parametryzacji rozwiązania oraz sposobu przetwarzania danych obrazowych (np. przekształcenia, normalizacja, rozszerzenie zbioru danych). Wyniki testów wydajnościowych i jakościowych dla wytrenowanych modeli.
4. Dokumentacja systemu oraz instrukcja wdrożenia systemu uwzględniająca niezbędne informacje dotyczące konfiguracji oprogramowania na nowych środowiskach oraz wyniki testów wydajnościowych.
5. Wdrożenie rozwiązania w ramach działalności Laboratorium promieniowania jonizującego GUM.

Faza III projektu będzie poświęcona na opublikowaniu wyniku prac badawczych, wdrożeniu najlepszych modeli i algorytmów uczenia maszynowego w ramach opracowanej finalnej wersji wybranej formy komercyjnej dystrybucji rozwiązania oraz integracji rozwiązania w ramach działalności operacyjnej GUM. Budżet realizacji fazy III zakłada pokrycie kosztów związanych z zatrudnieniem pracowników (min. naukowych i inżynierów), kosztów związanych z opracowaniem i publikacją wyników prac badawczych, wydzieleniem potrzebnych zasobów obliczeniowych oraz pokryciem kosztów pośrednich.