Specyfikacja Projektu

Wykorzystanie sztucznej inteligencji do opisania zdjęcia

## Opis dziedziny

Współczesne technologie, w tym modele językowe (LLM – Large Language Models), odgrywają coraz większą rolę w analizie i interpretacji obrazów. Systemy oparte na sztucznej inteligencji pozwalają na automatyczne generowanie opisów wizualnych w wielu dziedzinach życia, zarówno profesjonalnych takich jak agrokultura czy ochrona środowiska, jak i w życiu prywatnym do opisania zdjęć bądź wielu sytuacji. Systemy te wspierają użytkowników w analizie i interpretacji obrazów, co jest szczególnie przydatne dla osób nieposiadających specjalistycznej wiedzy w danym obszarze bądź też osób z niepełnosprawnością wzrokową.

W projekcie stworzymy system oparty na modelu językowym, który umożliwi generowanie opisów obrazów związanych z tematyką rolniczą. W ramach systemu model językowy przetworzy obrazy w formacie RGB, aby wygenerować ich ogólny opis. Obrazy te mogą przedstawiać takie elementy jak pola uprawne, rośliny czy problemy roślinne. System będzie narzędziem wspierającym analizę zawartości wizualnej, przydatnym w zarządzaniu uprawami.

## Wizja systemu

System ma umożliwić użytkownikom automatyczne generowanie ogólnych opisów obrazów, które koncentrują się na podstawowych elementach wizualnych. Opisy te będą krótkie i ogólne, co pozwoli na szybkie zrozumienie zawartości obrazu bez zagłębiania się w szczegóły. W związku z tym zadaniem modelu językowego jest wyłącznie faktograficzne opisanie obrazu.

**Główne funkcje systemu:**

1. **Przetwarzanie obrazów**: System przyjmuje jako dane wejściowe obrazy o stosunkowo małych rozmiarach, z których będą wyciągane kanały RGB. Obrazy rolnicze mogą przedstawiać różne elementy natury, np. pola, rośliny, szkodniki.
2. **Generowanie opisów**: Model językowy generuje opis obrazu, skupiający się na jego podstawowych elementach. Opisy są proste, zrozumiałe i niezbyt szczegółowe, aby były łatwe do interpretacji.
3. **Interfejs użytkownika**: Interfejs umożliwia użytkownikowi przesyłanie zdjęć i przeglądanie wygenerowanych opisów. Opis jest wyświetlany w formie tekstowej.

**Ograniczenia i założenia systemu:**

* Model generujący opis koncentruje się na podstawowych i ogólnych opisach bez dużej precyzji.
* System obsługuje tylko zdjęcia małego formatu, co pozwala na szybsze przetwarzanie.
* System dostarcza jedynie faktograficzny opis, bez zbędnej interpretacji.

**Cel systemu**: Celem projektu jest stworzenie narzędzia do automatycznego generowania ogólnych opisów obrazów, głównie związanych z agrokulturą. Dzięki temu użytkownik będzie mógł łatwo i szybko otrzymać informacje na temat obrazu, co okaże się przydatne w szczególności dla osób z niepełnosprawnością wzroku.