ALGORYTM ANONIMIZACJI

DOKUMENTACJA TECHNICZNA

Konrad Pusz
Weronika Kubińska
Anna Ratowska

1. WSTĘP

1.1 Wprowadzenie

Przy coraz szybszym rozwoju technologii, a także ciągłym zwiększaniu możliwości rejestrowania obrazu i wideo, problem prywatności stał się niezwykle ważnym zagadnieniem dla społeczeństwa. Samo utrwalanie wizerunku ludzi na zdjęciach często może nie być działaniem pożądanym, a jego dalsze rozpowszechnianie i udostępnianie bez wyraźnej zgody osoby przedstawionej na fotografii staje się już jednoznacznie czynem nielegalnym.

W raporcie omówiony został algorytm anonimizacji twarzy na zdjęciach. Jest to rozwiązanie wychodzące naprzeciw problemowi jednoznacznej identyfikacji osób widniejących na zdjęciu, bazujące na sztucznej inteligencji, a konkretniej na algorytmach uczenia maszynowego.

1.2 Anonimizacja twarzy a regulacje prawne

W wyniku wprowadzenia RODO w maju 2018 roku uregulowane zostało przetwarzanie danych osobowych. Dane osobowe to wszelkie informacje które mogą zidentyfikować osobę, na przykład określające jej fizyczną, lub kulturową tożsamość. Wizerunki osób na zdjęciach są danymi osobowymi, dlatego że pozwalają na ustalenie danych konkretnych osób. Podmioty publiczne i osoby fizyczne są zobowiązane do ochrony danych osobowych, w tym danych biometrycznych zawartych w obrazach i nagraniach wideo.

Dzięki algorytmowi anonimizacji możliwe jest zakrycie wizerunku twarzy na zdjęciu zgodnie z przepisami RODO. Celem jest skuteczne i nieodwracalne uniemożliwienie zidentyfikowania osoby fizycznej - ochrona prywatności przy jednoczesnym zachowaniu wiarygodności gromadzonych i wymienianych danych.

1.3. Ogólne cele anonimizacji

Praktyczne przykładowe zastosowania rozmycia twarzy i anonimizacji obejmują:

- ochronę prywatności i tożsamości w miejscach publicznych,
- ochronę osób w internecie (zamazywanie twarzy na przesyłanych zdjęciach),
- ochrona tożsamości dzieci,
- dziennikarstwo fotograficzne i raporty informacyjne (zamazywanie twarzy osób, które nie podpisały zgody),
- tworzenie i dystrybucja zbiorów danych (anonimizacja osób w zbiorach).

2. Wykorzystane narzędzia i biblioteki

W projekcie wykorzystane zostały różne biblioteki służące między innymi do przetwarzania obrazów i wykrywania obiektów - w tym celu zaimportowane zostały następujące pakiety:

- NumPy (Numerical Python) - biblioteka umożliwiająca zaawansowane obliczenia matematyczne, wspierająca szczególnie operacje na dużych macierzach. W programie wykorzystana do przetwarzania obrazów (obraz przedstawiony w postaci macierzy).

- OpenCV (Open Source Computer Vision) - biblioteka open source zapewniająca narzędzia i funkcje wykorzystywane w widzeniu komputerowym do przetwarzania i obróbki obrazów. Za jej pomocą można dokonać identyfikacji obiektów takich jak twarz na zdjęciach. Aby zidentyfikować wzór obrazu i jego różne cechy, używamy przestrzeni wektorowej i wykonujemy na tych cechach operacje matematyczne. Biblioteka posiada także zestaw funkcji służących do blurowania obrazów (rozmycie Gaussowskie, rozmycie uśredniające, rozmycie medianowe).

- imutils - pakiet bazujący na OpenCV, ułatwiający korzystanie z podstawowych funkcji przetwarzania obrazów takich jak translacja, obrót, zmiana rozmiaru, szkieletyzacja, wyświetlanie obrazów Matplotlib, sortowanie konturów czy wykrywanie krawędzi.

- dlib - biblioteka zawierająca algorytmy uczenia maszynowego i narzędzia do rozwiązywania rzeczywistych problemów. Posiada wiele narzędzi do rozpoznawania obiektów na obrazach, między innymi do rozpoznawania twarzy pod różnymi kątami z wysoką skutecznością. Implementuje takie algorytmy jak HOG (histogram ukierunkowanych gradientów) czy model CNN (konwolucyjna sieć neuronowa).

3. Działanie programu

3.1. Schemat działania programu

1) Utworzenie listy plików graficznych z podanej ścieżki, następnie wczytanie plików z listy.
2) Wybór jednego z dwóch trybów pracy algorytmu:
    a) Simple - używany jest model CNN (niezależnie od ilości wyznaczonych predykcji).
    b) Advanced - procedura anonimizacji przeprowadzana jest za pomocą tego modelu, który uzyskał lepszy wynik ilościowy w wykrywaniu twarzy na danym obrazie (CNN lub HOG).
3) Wykonanie działań zgodnie z wybranym trybem pracy
    a) Dla modelu HOG:
        i) Ekstrakcja wysokości i długości obrazu.
        ii) Konwersja obrazu z palety barw BGR do RGB.
        iii) Stworzenie kopii oryginalnego obrazu (by nie wykonywać wszystkich działań na oryginale).
        iv) Przetworzenie obrazu przez model HOG zainicjalizowany z biblioteki dlib.
    b) Dla modelu CNN:
        i) Przeskalowanie obrazu w celu optymalizacji pracy modelu CNN.
        ii) Konwersja obrazu z palety barw BGR do RGB.
        iii) Przetworzenie obrazu przez model CNN zainicjalizowany z biblioteki dlib.
        iv) Wyznaczenie koordynatów wszystkich ramek za pomocą przygotowanej funkcji.
        v) Stworzenie kopii oryginalnego obrazu, przygotowanie maski o rozmiarach obrazu.
    c) Po wykonaniu powyższych kroków, dla każdej twarzy znalezionej na zdjęciu (zarówno dla algorytmu HOG jak i CNN):
        i) Wyznaczenie/pobranie koordynatów startowych ramki oraz jej długości i szerokości.
        ii) Zablurowanie obszaru ramki na polu prostokąta za pomocą rozmycia uśredniającego (averaging blur) dla kopii obrazu.
        iii) Stworzenie bluru w kształcie elipsy dla maski, nałożenie kształtu maski na rozmyty obszar kopii obrazu, nałożenie maski na oryginalny obraz (finalne rozmycie w kształcie elipsy).
4) Zebranie wyników odnośnie ilości wykrytych twarzy na danym obrazie.
5) Wyświetlenie wyników - nazwy algorytmu, który dokonywał predykcji, ilości rozpoznanych twarzy oraz pliku wyjściowego z zablurowanymi twarzami.

3.2. Wykrywanie twarzy z użyciem HOG i CNN

W zależności od wybranego trybu pracy algorytmu, korzysta on z dwóch różnych wytrenowanych uprzednio modeli znajdujących się w bibliotece dlib - HOG face detector oraz CNN face detector. CNN uważany jest za dużo dokładniejszy detektor niż HOG, wymaga jednak więcej mocy obliczeniowej i zajmuje wyraźnie więcej czasu. W naszym programie dla trybu “Advanced” funkcja korzysta z tego z modeli, który dokonał lepszej detekcji.

3.2.1 HOG - Histogram of Oriented Gradients

HOG (Histogram of Oriented Gradients) - jest algorytmem wykrywania obiektów, twarzy. Ten deskryptor cech wykorzystuje się w technikach przetwarzania obrazów (image processing) oraz widzenia komputerowego (computer vision).

Na etapie przetwarzania wstępnego HOG bierze pod uwagę pięć filtrów:

1) przód twarzy,
2) twarz zwrócona w prawą stronę,
3) twarz zwrócona w lewą stronę,
4) przód twarzy, ale zwróconej w prawo,
5) przód twarzy, ale zwróconej w lewo.

Technika ta zlicza występowanie orientacji gradientu w określonej lokalizacji fragmentów obrazu. Koncentruje się ona na strukturze lub kształcie obiektu. Do obliczania cech wykorzystuje zarówno wielkość, jak i kąt gradientu. W skrócie w deskryptorze HOG jako cechy charakterystyczne używany jest rozkład (histogramy) kierunków gradientów (gradientów zorientowanych)

3.2.2 CNN - Convolutional Neural Network

CNN - Convolutional Neural Network - konwolucyjna sieć neuronowa to rodzaj sieci neuronowej z warstwami splotu.

W celu rozwiązania problemu rozpoznawania ludzkich twarzy na niewielkich oryginalnych zbiorach danych opracowano nowe podejście łączące konwolucyjną sieć neuronową (CNN) z powiększonym zbiorem danych. Oryginalny, mały zbiór danych jest powiększany do dużego zbioru poprzez szereg przekształceń obrazów twarzy. Na podstawie powiększonego zbioru danych obrazów twarzy można skutecznie wyodrębnić cechy twarzy i uzyskać wyższą dokładność rozpoznawania twarzy dzięki zastosowaniu sieci CNN.

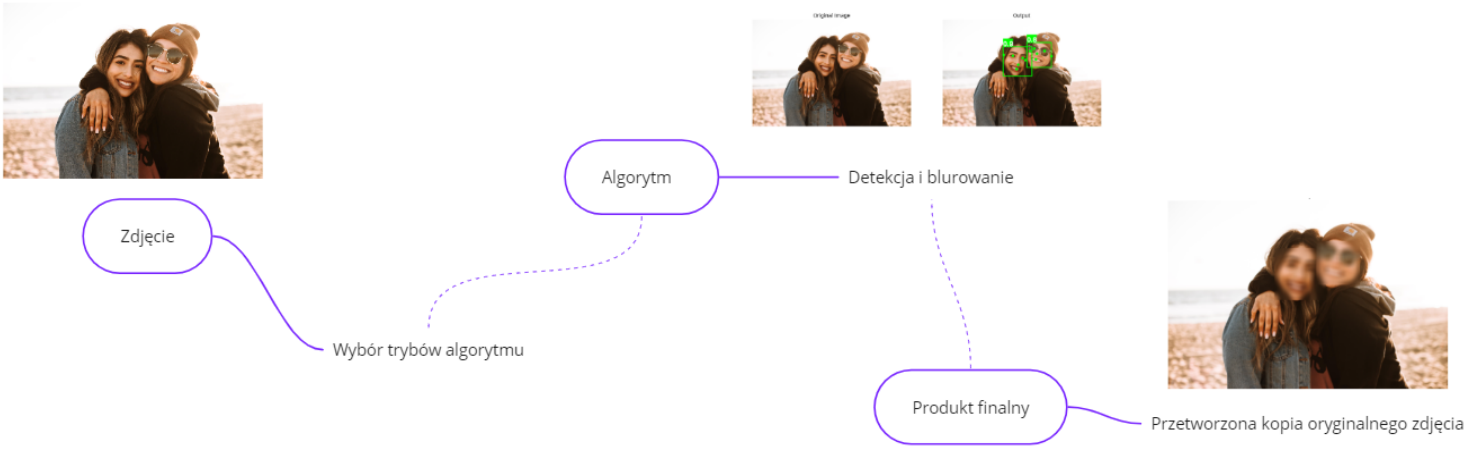
CNN zawiera dwa rodzaje warstw ukrytych, tj. warstwy konwolucyjne i warstwy łączące, które zazwyczaj są ułożone naprzemiennie w sieci neuronowej.

Wagi połączeń w sieci CNN mogą być współdzielone w całej sieci neuronowej, co pozwala nie tylko zmniejszyć ilość wag połączeń, ale także uprościć złożoność modelu sieci. Dzięki temu w większości przypadków można znacznie skrócić czas szkolenia sieci CNN. W szczególności, gdy obraz stanowi dane wejściowe sieci CNN, może on być wprowadzony bezpośrednio do sieci neuronowej, co pozwala uniknąć wielu skomplikowanych czynności, takich jak ekstrakcja cech i rekonstrukcja danych.

3.3 Blurowanie twarzy za pomocą rozmycia uśredniającego

Rozmycie (averaging blur) jest powszechnie stosowaną operacją przetwarzania obrazu służącą do redukcji szumu. Proces ten polega na usunięciu z obrazu zawartości o wysokiej częstotliwości, np. krawędzi, i wygładzeniu go.
Ogólnie rzecz biorąc, rozmycie obrazu uzyskuje się przez konwolucję obrazu przez jądro filtra dolnoprzepustowego. Podczas tej operacji obraz jest konwertowany za pomocą filtra pudełkowego (normalizowany). W tym procesie centralny element obrazu jest zastępowany średnią wszystkich pikseli w obszarze jądra (każdy element obrazu jest dodawany do swoich lokalnych sąsiadów, ważonych przez jądro). Przyjmuje się zasadę, że im większy rozmiar jądra, tym większy poziom rozmycia obrazu.

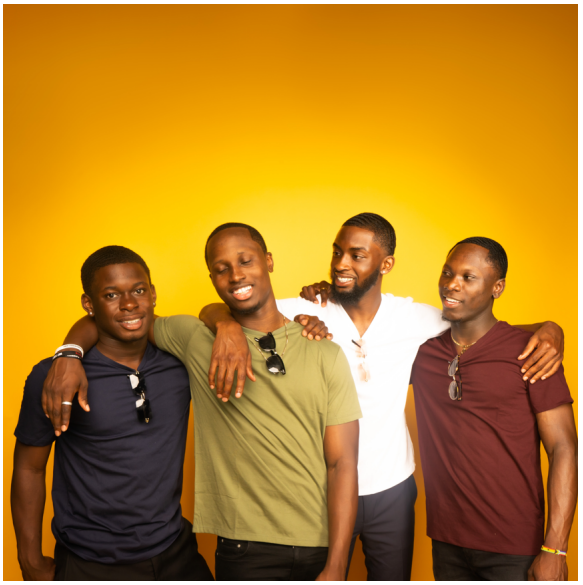
3.4 Graficzny schemat działania programu



3.5. Prezentacja działania programu na wybranym pliku graficznym

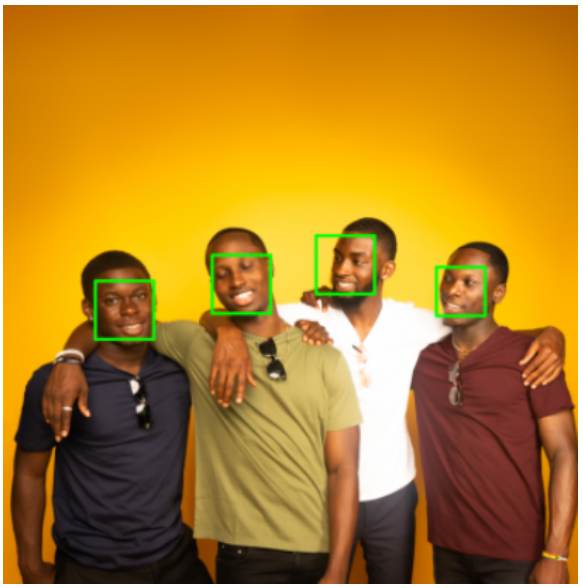
1) Oryginał zdjęcia dostarczany algorytmowi przez użytkownika.

Możemy zaobserwować, że fotografia przedstawia 4 osoby. Aby uniemożliwić identyfikację każdej z nich, chcemy wykorzystać nasz algorytm, który odpowiednio wykryje wszystkie twarze i dokona ich anonimizacji.



2) Wykrycie wszystkich twarzy na danym zdjęciu.

Zanim algorytm będzie mógł dokonać rozmycia twarzy, musi z jak największą dokładnością odnaleźć je na zdjęciu. Nasz program bardzo trafnie potrafi rozpoznawać twarze, nawet jeśli osoba ma zamknięte oczy lub jest zwrócona w bok. Za tę część odpowiadają modele HOG oraz CNN.



3) Finalny produkt anonimizacji

Ostatnim krokiem jest zablurowanie wszystkich znalezionych twarzy - jest to możliwe dzięki otrzymaniu dokładnych współrzędnych ramek twarzy zwróconych przez wspomniane wyżej modele. Przetworzenie tych obszarów za pomocą rozmycia uśredniającego powoduje, że każdy piksel zastępowany jest średnią innych pikseli w obszarze jądra, w efekcie czego osoby na zdjęciu stają się nierozpoznawalne.

4. Podsumowanie

RODO nie pozwala na przechowywanie zdjęć wykonanych bez zgody - anonimizacja danych to obecnie kluczowy element zgodności z przepisami.
Stworzony algorytm umożliwia zanonimizowanie fotografii dzięki technologii sztucznej inteligencji z wysoką skutecznością. Gromadzenie danych osobowych będzie maksymalnie ograniczone, ponieważ osoby, które mogłyby zostać zidentyfikowane, zostaną zamazane, aby nie można było ich już zidentyfikować.
Podsumowując, jest to niezwykle istotne dla prywatności i zgodności z dyrektywą, ponieważ w przypadku nieodwracalnej anonimizacji zdjęć nie ma ona zastosowania, gdyż nie dochodzi do przechowywania danych osobowych.

5. Kod źródłowy programu

Kod źródłowy programu znajduje się w notatniku Google Colab pod wskazanym linkiem:

https://colab.research.google.com/drive/1gXebtEWyNiTfUcsmHGvmvOQSkPNDZH4P?usp=sharing

Cały projekt dostępny również na repozytorium GitHub:

https://github.com/Conrad-Push/Automate_Face_Anonymization_Alghorithm

6. Bibliografia

- https://www.tutorialspoint.com/opencv/opencv_blur_averaging.htm
- https://pyimagesearch.com/2021/04/19/face-detection-with-dlib-hog-and-cnn/
- https://pyimagesearch.com/2021/04/28/opencv-smoothing-and-blurring/
- https://www.analyticsvidhya.com/blog/2022/04/face-detection-using-the-dlib-face-detector-model/
- https://www.tandfonline.com/doi/full/10.1080/21642583.2020.1836526
- https://towardsdatascience.com/cnn-based-face-detector-from-dlib-c3696195e01c