Spis treści

[1. Wykaz skrótów i symboli 2](#_Toc124464654)

[1.1. GPU 2](#_Toc124464655)

[1.2. CPU 2](#_Toc124464656)

[1.3. ∑ 2](#_Toc124464657)

[1.4. ∈ 2](#_Toc124464658)

[2. Wstęp 3](#_Toc124464659)

[3. Cel i zakres pracy 3](#_Toc124464660)

[4. Przegląd istniejących rozwiązań 4](#_Toc124464661)

[4.1. Prisma 4](#_Toc124464662)

[4.2. Deep Art Effects 4](#_Toc124464663)

[4.3. Adobe Photoshop 5](#_Toc124464664)

[4.4. Podsumowanie 6](#_Toc124464665)

[5. Zagadnienia 6](#_Toc124464666)

[5.1. Machine learning 6](#_Toc124464667)

[5.2. Sztuczna sieć neuronowa 7](#_Toc124464668)

[5.3. Deep neural network 7](#_Toc124464669)

[5.4. Convolutional neural network 7](#_Toc124464670)

[5.5. Loss function 7](#_Toc124464671)

[5.6. Neural representation 7](#_Toc124464672)

[5.7. Neural Style Transfer 7](#_Toc124464673)

[5.8. Content loss 7](#_Toc124464674)

[5.9. Style loss 7](#_Toc124464675)

[5.10. Total-variation loss 7](#_Toc124464676)

[6. Analiza wymagań funkcjonalnych 7](#_Toc124464677)

[6.1. Wczytanie obrazów 8](#_Toc124464678)

[6.2. Wyświetlanie obrazów 8](#_Toc124464679)

[6.3. Wczytanie modelu sieci 8](#_Toc124464680)

[6.4. Obliczanie utraty stylu i kontentu 8](#_Toc124464681)

[7. Wykorzystane technologie 8](#_Toc124464682)

[7.1. OpenCV 8](#_Toc124464683)

[7.2. .Net 7 8](#_Toc124464684)

[7.3. C# 8](#_Toc124464685)

[7.4. WPF 9](#_Toc124464686)

[8. Wykorzystane narzędzia programistyczne 9](#_Toc124464687)

[8.1. Git 9](#_Toc124464688)

[8.2. Visual Studio 2022 9](#_Toc124464689)

[9. Implementacja 10](#_Toc124464690)

[10. Podsumowanie 10](#_Toc124464691)

[Spis rysunków 10](#_Toc124464692)

[Bibliografia 10](#_Toc124464693)

# Wykaz skrótów i symboli

## GPU

GPU (ang. graphics processing unit), czyli procesor graficzny to wyspecjalizowany układ elektroniczny zaprojektowany do manipulowania i zmieniania pamięci w celu przyspieszenia tworzenia obrazów w buforze ramki przeznaczonym do wysłania do urządzenia wyświetlającego. Procesory graficzne są wykorzystywane w systemach wbudowanych, telefonach komórkowych, komputerach osobistych, stacjach roboczych i konsolach do gier. [1]

## CPU

CPU (ang. central processing unit), czyli jednostka centralna, zwana również procesorem centralnym, procesorem głównym lub po prostu procesorem, to układ elektroniczny wykonujący instrukcje składające się na program komputerowy. Procesor wykonuje podstawowe operacje arytmetyczne, logiczne, kontrolne oraz wejścia/wyjścia (I/O) określone przez instrukcje w programie. [2]

## ∑

Symbol ∑ oznacza sumowanie i jest używany jako skrócona notacja dla sumy wyrażeń, które są zgodne ze wzorem. [4]

## ∈

Symbol ∈ oznacza przynależność do zbioru i oznacza "należy do” więc stwierdzenie x∈A oznacza, że x jest elementem zbioru A. Innymi słowy, x jest jednym z obiektów w zbiorze (być może wielu) obiektów w zbiorze A. [5]

# Wstęp

Sztuczna inteligencja i głębokie uczenie to dwa obszary, które w ostatnich latach osiągnęły ogromny postęp. Dzięki rozwojowi algorytmów uczenia maszynowego i dostępności coraz większych zbiorów danych, sztuczna inteligencja jest w stanie rozwiązywać coraz bardziej skomplikowane zadania. W szczególności, głębokie uczenie pozwala na tworzenie sieci neuronowych, które są w stanie automatycznie uczyć się na podstawie danych i generować wyniki podobne do tych, które mogłyby zostać stworzone przez człowieka.

Jednym z ciekawszych zastosowań sztucznej inteligencji jest Neural Style Transfer - metoda pozwalająca na przeniesienie charakterystycznych cech jednego obrazu, takich jak kolorystyka, kompozycja czy tekstura, na inny obraz. Ta metoda jest często używana w grafice komputerowej, fotografii, reklamie czy sztuce. Przykładowo, może być używana do przeniesienia stylu malarskiego na zdjęcie lub stylizacja obrazu na stare zdjęcie.

Neural Style Transfer jest interesujący nie tylko ze względu na jego estetyczne zastosowanie, ale również ze względu na jego potencjał w rozwoju innych technologii. Metoda ta może być wykorzystywana w rozpoznawaniu i przetwarzaniu obrazów na podobnym poziomie, co człowiek. Neural Style Transfer może również pomóc w rozwoju nowych algorytmów, które pozwolą na automatyczne generowanie treści, takiej jak tworzenie filmów czy gier na podstawie istniejących obrazów.

Ten projekt pozwoli na przeanalizowanie możliwości Neural Style Transfer i jej potencjalnych zastosowań. Będzie to również okazja do poznania i zrozumienia różnych algorytmów i narzędzi, które są potrzebne do jego implementacji oraz zrozumienia jego działania. Praca jest przykładem jak sztuczna inteligencja i głębokie uczenie mogą być zastosowane w praktyce, tworząc innowacyjne rozwiązania dla różnych dziedzin.

# Cel i zakres pracy

Celem pracy jest zbudowanie aplikacji w technologii .NET, która umożliwi użytkownikom na łatwe i intuicyjne wykonanie transferu stylu na podstawie podanych przez nich zdjęć za pomocą technologii neural style transfer.

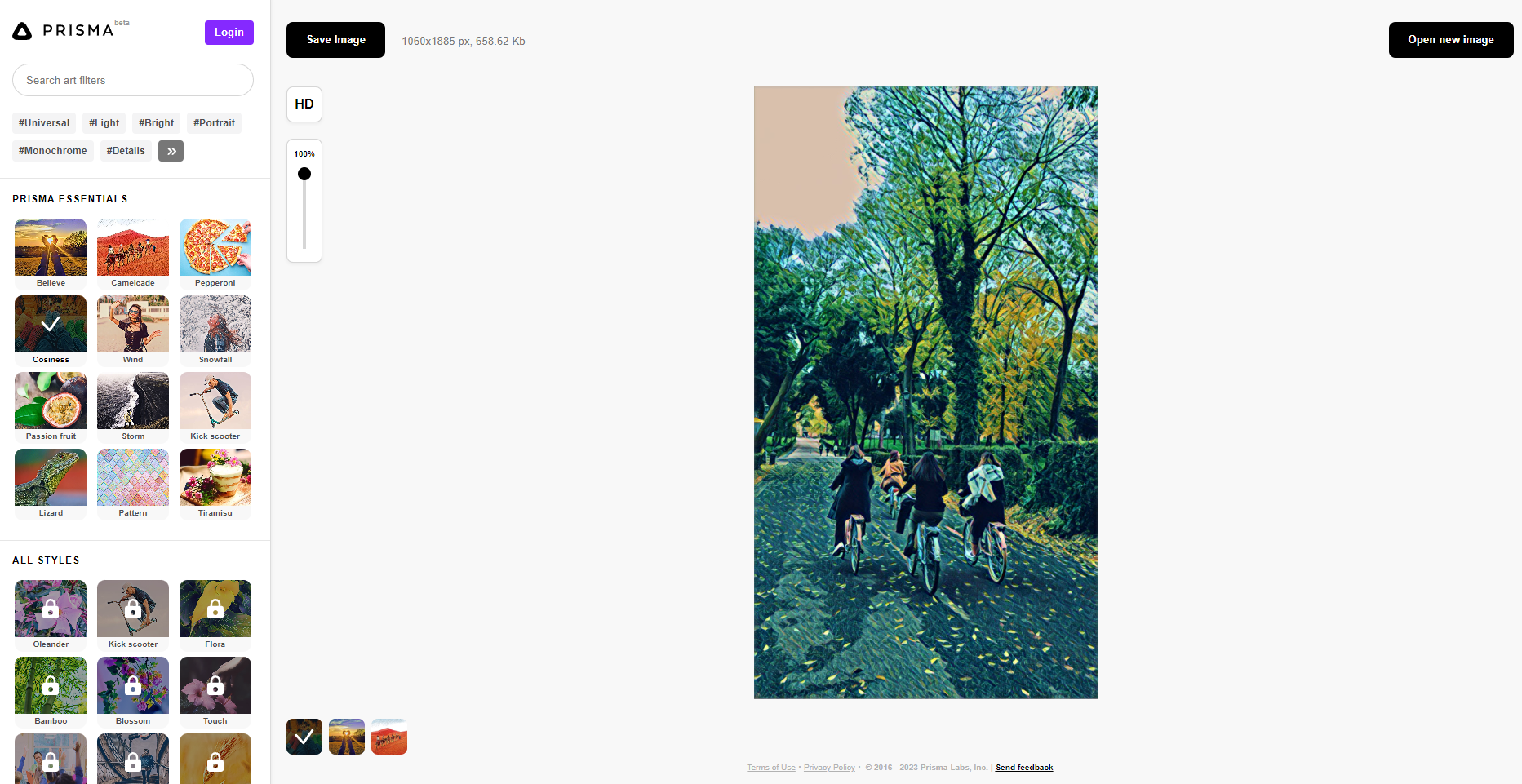
Zakres pracy (?)

# Przegląd istniejących rozwiązań

Obecnie, ze względu na dynamiczny rozwój rozwiązań opartych o sztuczną inteligencję, podobnych rozwiązań jest mnóstwo. Każdy program stara się być jak najbardziej dostępny dla użytkownika, łatwy w obsłudze i intuicyjny. Zebrane zostały trzy podobne rozwiązania, które różnią się od siebie sposobem działania, czy platformą na której są dostępne.

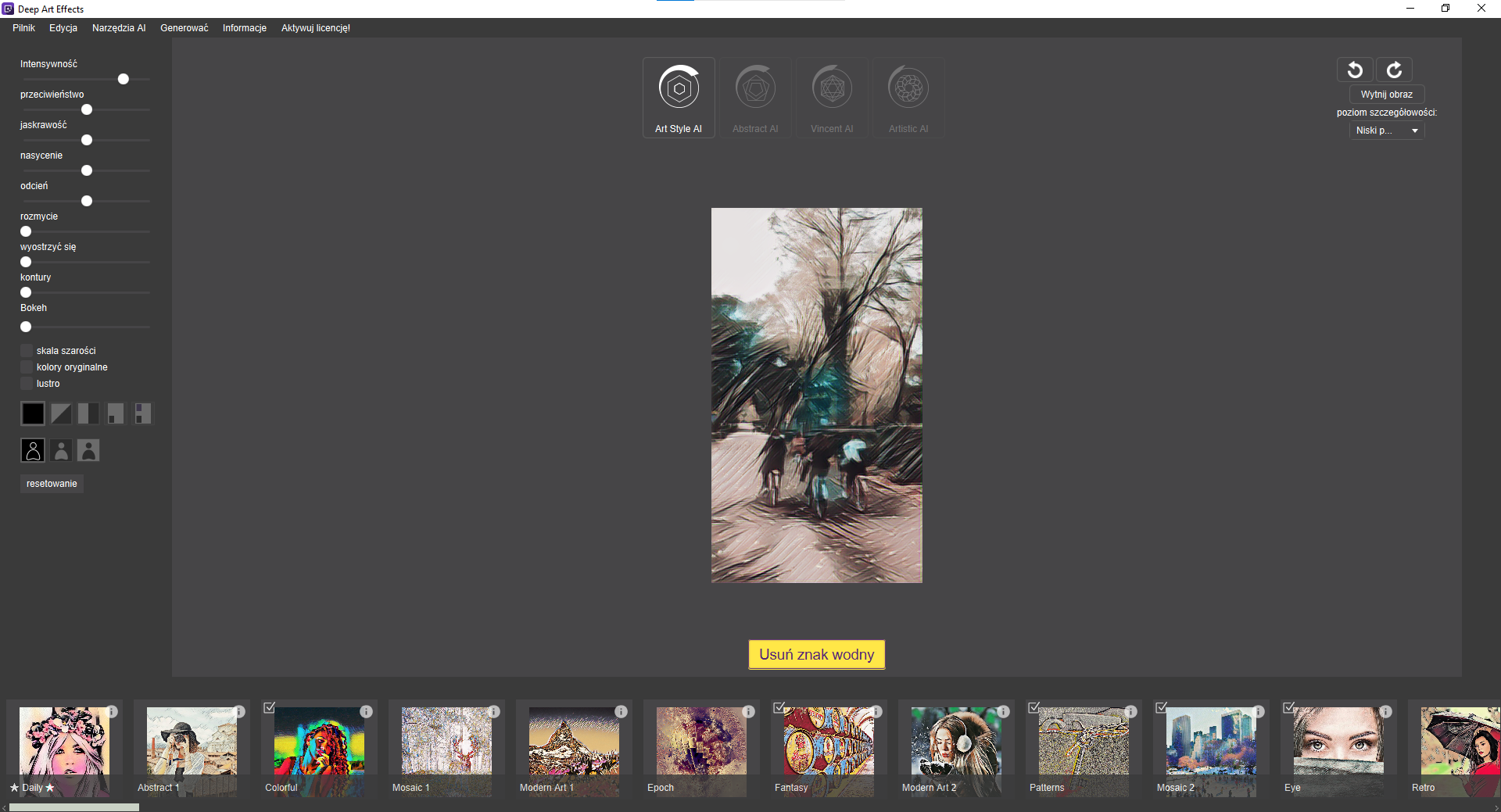
## Prisma

Prisma to aplikacja dostępna na urządzenia Android oraz iOS. Aplikacja pozwala na 3 dniowy okres próbny, a do kompletnego działania wymaga płatnego abonamentu. Na stronie producenta udostępniona jest także webowa wersja aplikacji, która za darmo umożliwia używanie tylko wybranych stylów. Bazowy obraz może zostać załadowny z urządzenia. Poniższy rysunek przedstawia przeglądarkową wersję aplikacji z załadowanym zdjęciem z komputera i wykorzystanym gotowym już obrazem stylu.



## Deep Art Effects

Aplikacja Deep Art Effects jest dostępna na urządzenia z systemem Android i iOS, oraz na komputery z systemem Windows. Tak jak Prisma, aplikacja jest płatna, ale umożliwia okres próbny.



Aplikacja na system Windows zajmuje 2GB pamięci na dysku oraz do prawidłowego działania wymaga 4GB pamięci operacyjnej RAM, jak twierdzi producent. Program potrafi zmieniać też inne parametry obrazu, takie jak kontrast, nasycenie i ma parę innych przydatnych funkcji.

## Adobe Photoshop

Adobe Photoshop oferuje funkcję generowania obrazu za pomocą technologii Neural Style Transfer. Korzystając z tej funkcji, użytkownicy mogą wybrać obraz źródłowy oraz obraz stylu, a następnie zastosować sztuczną inteligencję, aby przenieść styl obrazu stylu na obraz źródłowy. Możliwe jest także regulowanie poziomu szczegółowości oraz jasności, co jest istotne dla ostatecznego rezultatu. Użytkownicy mają także wybór, czy chcą zastosować tą technologię bezpośrednio z menu programu Adobe Photoshop, czy poprzez wtyczkę, co pozwala na uzyskanie bardziej zaawansowanych efektów. Technologia ta, dostępna w programie Adobe Photoshop jest uważana za jedną z najlepszych dostępnych na rynku, ze względu na bogactwo narzędzi, które pomagają w dalszej obróbce zdjęcia. Tak jak poprzednio wymienione aplikacje też jest płatny.

Obraz zawierający tekst, monitor, zrzut ekranu, ekran

Opis wygenerowany automatycznie

## Podsumowanie

Przedstawione zostały trzy różniące się od siebie aplikacje. Niestety każda ma swoje minusy. Wszystkie są płatne, ale posiadają okresy próbne. Jest to jedyny minus Prismy, pierwszej przedstawionej aplikacji, która w wersji przeglądarkowej jak i mobilnej, jest intuicyjna i szybka. Deep Art Effects, która dostępna jest w wersji desktopowej i mobilnej, ma duże wymagania sprzętowe i nie jest tak prosta w obsłudze jak Prisma. Adobe Photoshop jest potężną i bardzo popularną aplikacją, do której tak naprawdę możliwość transferu stylu została dodana jako opcja. Posiada mnóstwo możliwości, jednak dla użytkownika potrzebującego tylko zmienić styl zdjęcia będzie zbyt skomplikowana.

# Zagadnienia

## Machine learning

Machine learning, czyli uczenie maszynowe to dziedzina nauki, która pozwala komputerom na uczenie się bez konieczności jawnego programowania. [6]

W uczeniu maszynowym, algorytmy są trenowane, aby odkrywać wzorce i korelacje  
w dużych zbiorach danych, a następnie podejmować najlepsze decyzje i prognozy na podstawie tej analizy. W rezultacie, korzystanie z uczenia maszynowego pozwala na poprawę  
i zwiększenie dokładności danych, na których opierają się te decyzje. Zastosowania uczenia maszynowego są obecne wszędzie - w naszych domach, w sklepach, w mediach, w opiece zdrowotnej, w bankowości, w przemyśle, w bezpieczeństwie, w automatyzacji itd. [15]

Na podstawie przykładów znajdujących się w zbiorze uczącym, algorytmy uczenia maszynowego tworzą model matematyczny, który pozwala na prognozowanie lub podejmowanie decyzji bez konieczności jawnego programowania przez człowieka. Te algorytmy stosowane są w różnych dziedzinach, takich jak np. filtrowanie spamu lub rozpoznawanie obrazów, w których przypadku trudno jest opracować konwencjonalne algorytmy do wykonywania zadań. [7]

## Sztuczna sieć neuronowa

Sztuczna sieć neuronowa (ANN – ang. Adversary Neural Network), są ważnym elementem uczenia maszynowego i stanowią podstawę algorytmów uczenia głębokiego. Ich nazwa i struktura są inspirowane ludzkim mózgiem i naśladują sposób, w jaki neurony biologiczne komunikują się między sobą.

Sztuczne sieci neuronowe składają się z warstw węzłów, obejmujących warstwę wejściową, jedną lub więcej warstw ukrytych oraz warstwę wyjściową. Każdy węzeł połączony jest z innym i ma powiązaną wagę oraz próg. Jeśli wyjście pojedynczego węzła przekracza określoną wartość progową, węzeł ten jest aktywowany podczas przesyłania danych do kolejnej warstwy sieci. W przeciwnym razie żadne dane nie są przekazywane do następnej warstwy sieci. Sieci neuronowe uczy się na podstawie danych szkoleniowych, aby z upływem czasu poprawiać swoją dokładność. Gdy już ją osiągną, stają się potężnym narzędziem wykorzystywanym w informatyce i mechanizmach sztucznej inteligencji, które umożliwia błyskawiczne klasyfikowanie i podział danych na klastry. Zadania, takie jak rozpoznawanie mowy czy obrazów, trwają wtedy zaledwie kilka minut w porównaniu z godzinami identyfikacji przeprowadzanej odręcznie przez ekspertów. Jedną z najbardziej znanych sieci neuronowych jest algorytm wyszukiwania Google. [20]

## Deep neural network

Deep Neural Network (DNN) jest rodzajem sieci neuronowej, która charakteryzuje się wysokim poziomem złożoności, zwykle składającej się z co najmniej trzech warstw. DNN przetwarzają dane w złożony sposób przy użyciu zaawansowanych modeli matematycznych. Koncept DNN zaczął się rozwijać w latach 80-tych XX wieku, jednak dopiero ostatnie kilka lat przyniosło rewolucyjne zastosowanie tej technologii w różnych dziedzinach, takich jak rozpoznawanie obrazów, automatyzacja procesów biznesowych, analiza tekstu, czy naturalnego języka.

DNN posiada kilka charakterystycznych cech, które pozwalają na uzyskanie lepszych wyników niż w przypadku starszych modeli sieci neuronowych. Jednym z jej główych atutów jest możliwość radzenia sobie z nieuporządkowanymi danymi, co pozwala na uczenie sieci danymi, które nie zawierają wyraźnie określonych etykiet. Kolejną zaletą jest możliwość uczenia się hierarchii cech, co pozwala na rozwiązywanie złożonych problemów. Wreszcie, DNN cechuje się wysoką skutecznością, co pozwala na uzyskanie lepszych rezultatów niż w przypadku innych metod uczenia maszynowego.

Jednak jak każda technika , DNN również ma swoje wady. Jedną z głównych wad jest konieczność dużej ilości danych do trenowania sieci, co może być trudne do zdobycia dla niektórych problemów. Kolejną wadą jest potrzeba wysokiej mocy obliczeniowej do przetwarzania danych, co może stanowić barierę dla niektórych aplikacji. Wreszcie, ten model sieci cechuje się wysokim stopniem złożoności, co może utrudniać interpretację wyników i diagnozowanie problemów.

Pomimo tych wad, DNN staje się coraz bardziej popularne w wielu dziedzinach, zarówno ze względu na jego skuteczność, jak i rosnącą dostępność danych i mocy obliczeniowej. Obecnie DNN stosuje się w takich dziedzinach jak: medycyna, finanse, automatyka, rozpoznawanie mowy, czy obrazów. [21]

## Konwolucyjna Sieć Neuronowa

Konwolucyjna Sieć Neuronowa (CNN) to sieć neuronowa głębokiego uczenia, przeznaczona do przetwarzania uporządkowanych tablic danych. Jest szeroko stosowana w przetwarzaniu obrazów, rozpoznawaniu obiektów, rozpoznawaniu twarzy, detekcji ruchu i wykrywaniu obiektów. CNN również znalazło zastosowanie w przetwarzaniu języka naturalnego, takim jak klasyfikacja tekstu, czy tworzenie tłumaczeń maszynowych.

Sieci konwolucyjne są bardzo dobre w wykrywaniu wzorców w obrazie wejściowym, takich jak linie, gradienty, koła czy nawet oczy i twarze. To właśnie ta właściwość czyni je tak potężnymi w komputerowym przetwarzaniu obrazów. W przeciwieństwie do wcześniejszych algorytmów przetwarzania obrazów, sieci konwolucyjne mogą działać bezpośrednio na surowym obrazie.

Sieci te zawierają wiele warstw konwolucyjnych ułożonych jedna na drugiej, każda zdolna do rozpoznawania bardziej złożonych kształtów. Odzwierciedla to strukturę kory wzrokowej człowieka, gdzie seria warstw przetwarza obraz i identyfikuje coraz bardziej złożone cechy. Przy użyciu trzech lub czterech warstw konwolucyjnych możliwe jest rozpoznawanie ręcznie pisanych cyfr, a przy 25 warstwach możliwe jest rozróżnianie twarzy ludzkich.

Historia sieci konwolucyjnych sięga lat 80-tych, kiedy to pierwsze modele zaczęto budować na potrzeby rozpoznawania obrazów. Dzięki rozwojowi technologii i dostępności większej ilości danych, sieci konwolucyjne stały się bardziej zaawansowane i skuteczne, a ich zastosowanie rozszerzyło się na wiele dziedzin, takich jak medycyna czy tworzenie autonomicznych pojazdów. Sieci te są również stosowane w grach do analizowania ruchów gracza i przewidywania jego następnych ruchów. Stosuje się je także w wykrywaniu niestandardowych działań, takich jak kradzież, oszustwo czy niebezpieczne zachowanie. [22]

## Loss function

## Neural representation

## Neural Style Transfer

## Content loss

## Style loss

## Total-variation loss

# Analiza wymagań funkcjonalnych

W celu poprawnego wykonania aplikacji, konieczne jest wykonanie analizy założeń, jakie musi spełniać aplikacja.

## Wczytanie obrazów

## Wyświetlanie obrazów

## Wczytanie modelu sieci

## Obliczanie utraty stylu i kontentu

# Wykorzystane technologie

## OpenCV

OpenCV (ang. Open Source Computer Vision) to biblioteka programistyczna  
o otwartym kodzie źródłowym, która zapewnia narzędzia do przetwarzania obrazów i wideo. Pierwotnie została opracowana przez Gary'ego R. Bradsky w 1999 roku. Od tamtego czasu, biblioteka ta stała się jednym z najważniejszych narzędzi w dziedzinie analizy obrazów i wykrywania obiektów. OpenCV zawiera funkcje do operacji na obrazach, takie jak konwersja kolorów, filtrowanie, detekcja krawędzi, segmentacja, analiza obiektów i wiele innych. Biblioteka ta posiada także możliwość wczytywania i zapisywania modeli sieci neuronowych, co pozwala na wykorzystywanie modeli uczenia maszynowego w programach, które implementują tę bibliotekę. OpenCV jest również używana w aplikacjach mobilnych, takich jak aplikacje do rozpoznawania obrazów i analizy ruchu. Biblioteka ta jest stale rozwijana i aktualizowana przez społeczność programistów z całego świata, co pozwala na ciągły rozwój i poprawę jej funkcjonalności.

## .Net 6

.NET to platforma programistyczna firmy Microsoft, która pierwotnie została opracowana przez firmę w 2002 roku. Pierwsza wersja .NET Framework 1.0 została wydana wraz z systemem operacyjnym Windows XP. .NET Framework oferuje zestaw narzędzi i bibliotek, które umożliwiają programistom tworzenie aplikacji dla różnych systemów operacyjnych, takich jak Windows, Linux i macOS.

(…)

## C#

C# (ang. C-Sharp) to język programowania stworzony przez firmę Microsoft, który zadebiutował w 2000 roku jako jeden z języków programowania dla platformy .NET. Jest to obiektowy język wysokiego poziomu. C# jest językiem kompilowanym, co oznacza, że kod źródłowy jest kompilowany do kodu maszynowego przed uruchomieniem.

C# charakteryzuje się przejrzystą składnią, która jest zbliżona do C++ i Java, co ułatwia naukę języka. Jest to język typowo obiektowy z wsparciem dla wyrażeń lambda, wyjątków, interfejsów oraz dziedziczenia. Posiada także silne typowanie, co pozwala na szybkie wykrycie błędów w kodzie. Jest również integralną częścią środowiska .NET, co pozwala na wykorzystanie wszystkich narzędzi i bibliotek dostarczanych przez Microsoft. Jest wykorzystywany do tworzenia różnego rodzaju aplikacji, takich jak aplikacje okienkowe, aplikacje internetowe, gry, a także aplikacje mobilne oraz aplikacje IoT. W ciągu ostatnich lat C# stał się jednym z najważniejszych języków programowania w branży, dzięki swojej prostocie, wszechstronności i skalowalności. Warto też zwrócić uwagę na rosnące zastosowanie języka w tworzeniu aplikacji dla platform Unity, co pozwala na tworzenie gier i aplikacji VR. [9]

## WPF

Windows Presentation Foundation (WPF) jest platformą do tworzenia aplikacji dla systemu Windows, która została po raz pierwszy wprowadzona przez Microsoft w 2006 roku. Jest to podzbiór platformy .NET Framework, pozwala na tworzenie bogatych, interaktywnych i atrakcyjnych wizualnie aplikacji. WPF oferuje szeroki zestaw funkcji do tworzenia aplikacji.. Framework wykorzystuje język XAML (Extensible Application Markup Language), który zapewnia model deklaratywny do programowania aplikacji, co ułatwia programistom, którzy znają już .NET Framework za pomocą ASP.NET lub Windows Forms, rozpoczęcie pracy  
z WPF. Dodatkowo WPF pozwala na tworzenie aplikacji dla różnych urządzeń, takich jak tablety i smartfony, dzięki elastycznym i dostosowywalnym funkcjom. Framework posiada również różne efekty graficzne i animacje, co pozwala na bardziej atrakcyjne i interaktywne dla użytkownika aplikacje. [10]

## OpenCvSharp4.WpfExtensions

# Wykorzystane narzędzia programistyczne

## Git

Jest to system kontroli wersji, który umożliwia śledzenie zmian w kodzie. Dzięki temu możliwe jest przeglądanie historii modyfikacji poszczególnych fragmentów kodu, a nawet całego projektu. Git udostępnia narzędzia do rozgałęzionego rozwoju kodu poprzez tworzenie oddzielnych gałęzi, które mogą rozszerzać wybrane, inne gałęzie. Kod jest przechowywany w zewnętrznym repozytorium, co zabezpiecza go przed np. awarią sprzętu.

## Visual Studio 2022

Visual Studio 2022 to zintegrowane środowisko programistyczne (IDE) dostarczane przez Microsoft, przeznaczone do tworzenia aplikacji dla różnych platform, takich jak Windows, Web, Cloud, iOS i Android. Zawiera narzędzia do tworzenia, debugowania, testowania i profilowania kodu, a także narzędzia do zarządzania projektem, współpracy z zespołem i integracji z systemami kontroli wersji. Visual Studio 2022 oferuje wsparcie dla wielu języków programowania, w tym C#, C++, F#, JavaScript, Python i innych.

# Implementacja

# Podsumowanie

# Spis rysunków

# Bibliografia

[1] “Graphics processing unit” <https://en.wikipedia.org/wiki/Graphics_processing_unit#cite_note-1>

[2] “Central processing unit” <https://en.wikipedia.org/wiki/Central_processing_unit>

[3]

[4] Definition of the summation symbol.” From Math Insight. <http://mathinsight.org/definition/summation_symbol>

[5] Nykamp DQ, “Definition of the set membership symbol.” From Math Insight. <http://mathinsight.org/definition/set_membership_symbol>

[6] Machine Learning, Tom Mitchell, McGraw Hill, 1997.

[7] <https://pl.wikipedia.org/wiki/Uczenie_maszynowe>

[20] https://www.ibm.com/pl-pl/cloud/learn/neural-networks

[15] https://www.sap.com/poland/insights/what-is-machine-learning.html

[9] “A tour of the C# language” https://learn.microsoft.com/en-us/dotnet/csharp/tour-of-csharp/

[10] “Get started (WPF)” <https://learn.microsoft.com/en-us/dotnet/desktop/wpf/getting-started/>

[21] <https://www.techopedia.com/definition/32902/deep-neural-network>

[22] https://deepai.org/machine-learning-glossary-and-terms/convolutional-neural-network