Хмельницький національний університет
Факультет інформаційних технологій
Кафедра комп'ютерних наук

КВАЛІФІКАЦІЙНА РОБОТА МАГІСТРА

на тему *Метод виявлення зовнішніх проявів насильства у відеопотоці нейромережевими засобами*

Галузь знань *12 – Інформаційні технології*
Шифр і назва галузі знань

Спеціальність *122 – Комп'ютерні науки*
Шифр і назва спеціальності

Освітня програма *Комп'ютерні науки*
Назва освітньої програми

Виконав: *студент 2 курсу, група КНм-22-1* _____ *Е.Р. Муляр*
Курс, група виконавця | Підпис | Ініціали, прізвище

Керівник: *к.т.н., доцент кафедри КН* _____ *Р.О. Багрій*
Науковий ступінь, посада | Підпис | Ініціали, прізвище

Нормоконтроль: *к.т.н., доцент кафедри КН* _____ *Р.О. Багрій*
Науковий ступінь, посада | Підпис | Ініціали, прізвище

До захисту допускаю:
Зав. кафедри КН, д.т.н., професор _____ *О.В. Бармак*
Підпис | Ініціали, прізвище

04 грудня 2023 р.

Хмельницький 2023

ХМЕЛЬНИЦЬКИЙ НАЦІОНАЛЬНИЙ УНІВЕРСИТЕТ

Факультет *інформаційних технологій*

Кафедра *комп'ютерних наук*

Освітній ступінь *магістр*

Галузь знань *12 – Інформаційні технології*

Спеціальність *122 – Комп'ютерні науки*

ЗАТВЕРДЖУЮ

Завідувач кафедри *комп'ютерних наук*

(підпис)

д.т.н., професор О.В. Бармак

«*01*» *вересня* *2023* року

ЗАВДАННЯ

НА КВАЛІФІКАЦІЙНУ РОБОТУ МАГІСТРА

1. Тема кваліфікаційної роботи магістра: *«Метод виявлення зовнішніх проявів насильства у відеопотоці нейромережевими засобами»*

2. Завдання видано студенту *Муляру Едуарду Руслановичу*
(прізвище, ім'я, по батькові)

3. Керівник роботи *доцент кафедри КН Багрій Руслан Олександрович*
(прізвище, ім'я, по батькові)

4. Затверджені наказом університету від «*15*» *серпня* *2023* р. № *30*

5. Зміст пояснювальної записки (перелік задач) та вихідні дані:

Мета кваліфікаційної роботи магістра – розробка методу виявлення зовнішніх проявів насильства у відеопотоці нейромережевими засобами. Для досягнення поставленої мети визначено наступні задачі: провести аналіз нейромережевих моделей та існуючих підходів для виявлення проявів насильства у відеопотоці; розробити метод виявлення зовнішніх проявів насильства у відеопотоці з використанням згорткової нейронної мережі та класифікатора SVM; підготувати набір даних для навчання згорткової нейронної мережі; навчити попередньо навчену згорткову нейронну мережу виявляти ознаки насильства на неперервному відеопотоці даних; визначити загальну точність запропонованого методу виявлення зовнішніх проявів насильства.

Реферат

Кваліфікаційна робота магістра присвячена розробці методу виявлення зовнішніх проявів насильства у відеопотоці нейромережевими засобами.

Актуальність теми. Насильство є однією з найбільш серйозних проблем сучасного суспільства. Часто злочинні дії фіксуються на відео, але їх розпізнавання та ідентифікація залишається складним завданням для правоохоронних органів.

Метод виявлення зовнішніх проявів насильства у відеопотоці нейромережевими засобами є актуальною темою дослідження. Згорткові нейронні мережі є ефективним інструментом для аналізу зображень та відео, і використовуються для розпізнавання обличь, предметів та образів. Метод виявлення насильства у відеопотоці з використанням нейронних мереж полягає у використанні алгоритмів машинного навчання для виявлення певних ознак насильства на відео. Ці ознаки можуть включати гучні крики, рухи, які свідчать про фізичне насильство, та інші ознаки [1]. Однією з переваг цього методу є можливість автоматичного виявлення насильства на великій кількості відео, що дозволяє економити час та зусилля людей, які займаються цією проблемою. Крім того, цей метод може бути використаний для виявлення насильства на «живому відео», що дозволяє більш оперативно реагувати на події.

Отже, метод виявлення зовнішніх проявів насильства у відеопотоці нейромережевими засобами є перспективною темою дослідження, яка може сприяти зменшенню насильства в суспільстві та забезпечити більш ефективну боротьбу з цією проблемою.

Мета і задачі роботи. Метою кваліфікаційної роботи магістра є розробка методу виявлення зовнішніх проявів насильства у відеопотоці нейромережевими засобами.

Для досягнення поставленої мети визначено наступні задачі:

– провести аналіз нейромережевих моделей та існуючих підходів для виявлення проявів насильства у відеопотоці;

– розробити метод виявлення зовнішніх проявів насильства у відеопотоці з використанням згорткової нейронної мережі та класифікатора SVM;

– підготувати набір даних для навчання згорткової нейронної мережі;

– донавчити попередньо навчену згорткову нейронну мережу виявляти ознаки насильства на неперервному відеопотоці даних;

– провести валідацію запропонованого методу виявлення зовнішніх проявів насильства за стандартними статистичними показниками.

Об'єкт дослідження – процес виявлення прояву насильства у відеопотоці нейромережевими засобами.

Предмет дослідження – моделі нейронної мережі, методи класифікації ознак для виявлення прояву насильства у відеопотоці.

Методи дослідження, застосовані для вирішення поставлених завдань: для виявлення ознак насильства на відео – згорткова нейронна мережа; для класифікації проявів насильства – метод опорних векторів.

Наукова новизна одержаних результатів. В результаті проведеної роботи були отримані наступні результати:

– вдосконалено архітектуру згорткової нейронної мережі, що дало можливість виявляти прояви насильства у відеопотоці даних у реальному часі, що досягається за рахунок додаткового навчання попередньо навченої моделі на неперервному відеопотоці даних який містить прояви насильства;

– розроблено метод виявлення зовнішніх проявів насильства у відеопотоці за допомогою згорткової нейронної мережі та класифікатора SVM, що дозволило підвищити точність виявлення проявів насильства до 87.4%-99.45% у неперервному відеопотоці у реальному часі.

Апробація результатів кваліфікаційної роботи магістра та публікації. Основні положення і результати роботи опубліковані в збірнику наукових праць – Метод виявлення ознак насильства у відеоматеріалах нейромережевими засобами / Муляр Е.Р., Багрій Р.О., Манзюк Е.А., Пасічник О.А. // Збірник наукових праць за

матеріалами Всеукраїнської науково-практичної конференції «Актуальні проблеми комп’ютерних наук – 2023» Хмельницький, 2023.

Також, результати роботи опубліковані у науковому журналі – Метод виявлення зовнішніх проявів насильства у відеопотоці нейромережевими засобами / Муляр Е.Р., Багрій Р.О., Манзюк Е.А., Пасічник О.А. // Науковий журнал «Вісник Хмельницького національного університету. Серія: Технічні науки» Хмельницький, 2023, №6 (Довідка з редакції).

Структура та обсяг роботи. Кваліфікаційна робота магістра складається з завдання, реферату, змісту, переліку скорочень, вступу, 4 розділів, висновків, переліку посилань із 44 найменувань та 4 додатків. Загальний обсяг кваліфікаційної роботи магістра становить 77 сторінок, з них 71 сторінок основного тексту та 30 сторінок додатків. У роботі наведено 30 рисунків.

Ключові слова: насильство, виявлення, відеопотік, реальний час, нейромережі, згорткова нейронна мережа, SVM.

Зміст

Перелік скорочень

Скорочення, термін, позначення	Пояснення
IT	Інформаційні технології
ШНМ	Штучна нейронна мережа
LSTM	Long short-term memory
RNN	Recurrent neural networks
CNN	Convolutional neural network
SVM	Support vector machines

Вступ

Актуальність теми. Насильство є однією з найбільш серйозних проблем сучасного суспільства. Часто злочинні дії фіксуються на відео, але їх розпізнавання та ідентифікація залишається складним завданням для правоохоронних органів.

Метод виявлення насильства на відео за допомогою згорткової нейронної мережі є актуальною темою дослідження. Згорткові нейронні мережі є ефективним інструментом для аналізу зображень та відео, і використовуються для розпізнавання обличь, предметів та образів. Метод виявлення насильства на відео з використанням згорткової нейронної мережі полягає у використанні алгоритмів машинного навчання для виявлення певних ознак насильства на відео. Ці ознаки можуть включати гучні крики, рухи, які свідчать про фізичне насильство, та інші ознаки [1]. Однією з переваг цього методу є можливість автоматичного виявлення насильства на великій кількості відео, що дозволяє економити час та зусилля людей, які займаються цією проблемою. Крім того, цей метод може бути використаний для виявлення насильства на «живому відео», що дозволяє більш оперативно реагувати на події.

Отже, метод виявлення насильства на відео за допомогою згорткової нейронної мережі є перспективною темою дослідження, яка може сприяти зменшенню насильства в суспільстві та забезпечити більш ефективну боротьбу з цією проблемою.

Мета і задачі роботи. Метою кваліфікаційної роботи магістра є розробка методу виявлення зовнішніх проявів насильства у відеопотоці нейромережевими засобами.

Для досягнення поставленої мети визначено наступні задачі:

– провести аналіз нейромережевих моделей та існуючих підходів для виявлення проявів насильства у відеопотоці;

– розробити метод виявлення зовнішніх проявів насильства у відеопотоці з використанням згорткової нейронної мережі та класифікатора SVM;

– підготувати набір даних для навчання згорткової нейронної мережі;

– донавчити попередньо навчену згорткову нейронну мережу виявляти ознаки насильства на неперервному відеопотоці даних;

– провести валідацію та аналіз отриманих експериментальних результатів запропонованого методу виявлення зовнішніх проявів насильства за стандартними статистичними показниками.

Об’єкт дослідження – процес виявлення прояву насильства у відеопотоці нейромережевими засобами.

Предмет дослідження – моделі нейронної мережі, методи класифікації ознак для виявлення прояву насильства у відеопотоці.

Методи дослідження, застосовані для вирішення поставлених завдань: для виявлення ознак насильства на відео – згорткова нейронна мережа; для класифікації проявів насильства – метод опорних векторів.

Наукова новизна одержаних результатів. В результаті проведеної роботи були отримані наступні результати:

– вдосконалено архітектуру згорткової нейронної мережі, що дало можливість виявляти прояви насильства у відеопотоці даних у реальному часі, що досягається за рахунок додаткового навчання попередньо навченої моделі на неперервному відеопотоці даних з проявами насильства;

– розроблено метод виявлення зовнішніх проявів насильства у відеопотоці за допомогою згорткової нейронної мережі та класифікатора SVM, що дозволило підвищити точність виявлення проявів насильства до 87.4%-99.45% у неперервному відеопотоці у реальному часі.

Апробація результатів кваліфікаційної роботи магістра та публікації. Основні положення і результати роботи опубліковані в збірнику наукових праць – Метод виявлення ознак насильства у відеоматеріалах нейромережевими засобами / Муляр Е.Р., Багрій Р.О., Манзюк Е.А., Пасічник О.А. // Збірник наукових праць за матеріалами Всеукраїнської науково-практичної конференції «Актуальні проблеми комп’ютерних наук – 2023» Хмельницький, 2023.

Також, результати роботи опубліковані у науковому журналі – Метод виявлення зовнішніх проявів насильства у відеопотоці нейромережевими засобами / Муляр Е.Р., Багрій Р.О., Манзюк Е.А., Пасічник О.А. // Науковий журнал «Вісник Хмельницького національного університету. Серія: Технічні науки» Хмельницький, 2023, №6 (Довідка з редакції).

Структура та обсяг роботи. Кваліфікаційна робота магістра складається з завдання, реферату, змісту, переліку скорочень, вступу, 4 розділів, висновків, переліку посилань із 44 найменувань та 4 додатків. Загальний обсяг кваліфікаційної роботи магістра становить 77 сторінок, з них 71 сторінок основного тексту та 30 сторінок додатків. У роботі наведено 30 рисунків.

РОЗДІЛ 1
Аналіз сучасного стану використання інформаційних технологій для виявлення насильства на відео

1.1. Аналіз предметної області

У наш час насильство є серйозною проблемою, яка поширюється і має негативний вплив на суспільство. Вплив насильства можна побачити в різних формах людського побуту у всіх частинах світу. Щороку більше мільйона людей втрачають життя, і набагато більше людей отримують несмертельні травми внаслідок насильства, спричиненого самим собою, міжособистісного чи колективного насильства. Загалом, насильство є однією з головних причин смерті людей у віці 15-44 років у всьому світі [2].

Насильство проявляється у різних формах:

– *психологічне* – вид насильства, яке включає у себе вербальні нападки, погрози, приниження, переслідування, страхування та інші дії, спрямовані на обмеження волі людини, контроль над її репродуктивною поведінкою, а також спричинити емоційну невпевненість, нездатність захистити себе або завдати шкоди її психічному здоров'ю;

– *фізичне* – вид насильства, що являє собою удари, штовхання, захоплення, биття, кусання, а також незаконне позбавлення волі, заподіяння тілесних ушкоджень різної тяжкості, залишення в небезпеці, відмова в допомозі людині, яка перебуває в небезпечному стані для життя, заподіяння смерті та інші подібні дії;

– *економічне* – вид насильства, яке включає в себе цілеспрямоване позбавлення особи житла, харчів, одягу, майна, коштів, документів або можливості користуватися ними, а також бездоглядність або не надання необхідної допомоги в лікуванні, реабілітації чи інших послугах, заборона працювати або навчатися, примушення до праці та інші порушення права економічного характеру [3];

– *сексуальне* – вид насильства, яке включає будь-які сексуальні дії, здійснені проти дорослої особи без її згоди або проти дитини незалежно від її згоди, або в

присутності дитини; також може включати примушування до сексуальних актів з третьою особою і будь-яке інше порушення статевої свободи або статевої недоторканності, зокрема вчинене щодо дитини або в її присутності [4].

У сучасному світі для боротьби з насильством використовують системи відеоспостережень. Відеоспостереження є цінним інструментом для моніторингу людської поведінки. Камери відеоспостереження встановлюються в громадських і приватних місцях, таких як лікарні, школи та в'язниці, щоб виявляти та запобігати насильницькій поведінці [5].

Камери відеоспостереження є високоефективним рішенням для забезпечення безпеки людей. Відеокамери використовуються для дистанційного спостереження за складними середовищами в різних погодних умовах, що включають постійну зміну освітленості, температури та видимості. Виробники обладнання для відеоспостереження вирішують ці проблеми, надаючи сучасні камери, які мають більш високу роздільну здатність, вбудовані нагрівальні елементи, автоматичне покращення зображення в умовах слабкого освітлення, а також технології кодування та шифрування.

Існує декілька видів камер відеоспостереження, які можуть бути використані для виявлення насильства на відео:

– *аналогові камери*, які відправляють сигнал напряму до монітора або відеореєстратора. Вони прості в установці і використанні, але їхні можливості щодо виявлення насильства обмежені;

– *IP-камери*, підключаються до мережі і передають відео та аудіосигнали через Інтернет; вони можуть мати вбудовані функції виявлення руху та інші алгоритми аналізу, що поліпшують ефективність виявлення насильства: сучасні камери з Інтернет-протоколом мають роздільну здатність 4K, яка в 27 разів перевищує роздільну здатність звичайних аналогових камер;

– *PTZ-камери*, мають можливість панорамного огляду, нахилу та зуму; вони можуть відстежувати рух об'єктів і автоматично фокусуватися на потенційних джерелах насильства, що покращує їхню ефективність [6].

Ефективність камер відеоспостереження у виявленні насильства на відео залежить від декількох факторів, включаючи:

– висока роздільна здатність камери дозволяє отримати деталізоване зображення, що поліпшує здатність виявлення насильства і ідентифікацію осіб;

– камери з високою чутливістю до освітлення можуть ефективно працювати як у денний, так і у нічний час, забезпечуючи якісну зйомку навіть за складних умов освітлення;

– камери з вбудованими функціями аналізу відео, такими як виявлення руху, аналіз поведінки людей та інші алгоритми, дозволяють автоматично виявляти можливі випадки насильства і сповіщати оператора [6].

Основною проблемою процесу відеоспостереження є людський фактор: неуважність та недбалість. Людина-оператор може ефективно стежити лише за кількома камерами відеоспостереження, але часто буває обтяжений великою кількістю камер, що призводить до помилок і пропусків виявлення. Для того, щоб усунути цю проблему людство застосовує інформаційні технології. Сьогодні найефективнішим інструментом для виявлення насильства під час відеоспостереження є штучні нейронні мережі.

Штучна нейронна мережа є математичною моделлю, що імітує роботу людського мозку. Вона складається з взаємопов'язаних штучних нейронів, які передають інформацію один одному через зв'язки з різною вагою. ШНМ призначена для вирішення певних завдань, таких як класифікація зображень, розпізнавання мови, прогнозування та інші [7].

Використання ШНМ у відеоспостереженні є ефективним інструментом для автоматизації процесу виявлення насильства.

Даний підхід має декілька переваг:

– ШНМ можуть виявляти насильство на відео з високою точністю, оскільки вони можуть аналізувати велику кількість даних та виявляти навіть тіньові ознаки насильства;

– ШНМ можуть працювати швидко та ефективно, що дозволяє їм аналізувати великі обсяги відео за короткий час;

– ШНМ можуть працювати без участі людей, що дозволяє автоматизувати процес виявлення насильства на відео та зменшує витрати на персонал;

– ШНМ можуть навчатися на нових даних та покращувати свою точність виявлення насильства на відео з часом.

ШНМ можуть використовуватися в різних галузях, таких як медіа, безпека, правоохоронні органи та інші, що робить їх універсальними засобами виявлення насильства на відео [8].

Найпоширенішим на сьогодні видом нейронних мереж, який використовується в області виявлення ознак насильства у відеопотоці є нейронні мережі, які базуються на глибокому навчанні [9].

Глибинне навчання є спеціалізованою формою машинного навчання, в якій робочий процес починається з вилучення відповідних характеристик із зображень. Замість ручного вилучення, глибинне навчання використовує автоматичний підхід, щоб витягнути ці характеристики зображень. У глибинному навчанні також застосовується «наскрізне навчання», тобто мережі надаються вихідні дані та відповідне завдання, для прикладу класифікація, і мережа автоматично навчається, як це зробити. Це означає, що глибинна нейронна мережа сама визначає та витягує характеристики зображень, що необхідні для класифікації об'єктів на зображенні.

Алгоритми глибинного навчання можуть масштабуватися зі зростанням даних, тоді як неглибинне навчання досягає певного рівня продуктивності і потім сходиться. У неглибинному навчанні, коли до мережі додаються більше прикладів та навчальних даних, досягається певний рівень продуктивності. Проте, основною перевагою глибинного навчання є те, що ці мережі зазвичай вдосконалюються зі збільшенням обсягу доступних даних. У машинному навчанні, ознаки та класифікатори для сортування зображень вибираються вручну. Однак у глибинному навчанні ці етапи виконуються автоматично, завдяки здатності моделей глибинного навчання самостійно виокремлювати ознаки та розробляти моделі [10].

Для навчання глибинної мережі можна використовувати різні типи навчання, такі як навчання з учителем, яке передбачає наявність множини тренувальних анотованих даних, або навчання без учителя. В глибинному навчанні існують три типи шарів нейронів у нейронній мережі: вхідний шар, приховані шари і вихідний шар. Зв'язки між нейронами мають вагу, яка визначає важливість вхідних даних. Глибинна нейронна мережа є мережею, яка має більше двох прихованих шарів. Для навчання глибинної нейронної мережі є необхідним мати значний обсяг тренувальних даних. Чим більший набір даних, тим краще можна навчити глибинну мережу виконувати завдання класифікації або прогнозування.

Нейрони у глибинних мережах організовані у три типи шарів:

– *вхідний шар* отримує вхідні дані і передає їх до першого прихованого шару – цей шар відповідає за приймання вхідних сигналів та передачу їх далі у мережу;

– *прихований шар* виконує математичні обчислення на вхідних даних – ці шари відповідають за обробку та аналіз вхідних сигналів, виконуючи різні обчислювальні операції; важливим аспектом при створенні таких мереж є визначення кількості прихованих шарів та кількості нейронів у кожному з них;

– *вихідний шар* повертає оброблені дані в якості вихідного результату мережі – цей шар відповідає за генерацію вихідного сигналу або прогностичних значень на основі оброблених даних [11].

Така організація нейронів у шари допомагає глибинним мережам виконувати складні обчислення та здійснювати класифікацію, прогнозування або інші завдання, залежно від призначення мережі.

Для досягнення задовільного рівня точності програми глибинного навчання необхідно мати доступ до величезних обсягів навчальних даних та потужних обчислювальних ресурсів. Оскільки програми глибинного навчання можуть створювати складні статистичні моделі, базуючись на своїх власних ітераційних результатів, вони здатні створювати точні прогностичні моделі, використовуючи великі обсяги немаркованих та неструктурованих даних. Це особливо важливо у

зв'язку зі зростаючою популярністю Інтернету речей (IoT), оскільки більшість даних, які створюються як людьми, так і машинами, є немаркованими та неструктурованими.

Порівнюючи послідовно отримані результати з включеними в набір даними, можна виконати ітерації та обчислити функцію втрат, яка вказує на ступінь помилок алгоритму. Після кожної ітерації, так званої епохи, ваги між нейронами перерозподіляються за допомогою методу градієнтного спуску, що дозволяє мінімізувати значення функції втрат.

Існує кілька способів використання нейронних мереж, зокрема:

– тренування штучної нейронної мережі на власних даних – цей підхід передбачає навчання мережі на власному наборі даних, де алгоритм самостійно встановлює ваги нейронів, щоб відповідати задачі, яку потрібно вирішити;

– використання готової архітектури штучної нейронної мережі з попередньо навченими вагами – цей підхід включає використання вже натренованої моделі, яка була підготовлена на певному наборі даних для певної кількості класів; це особливо корисно, коли у нас обмежені обчислювальні ресурси або недостатньо власних даних для тренування;

– дотренування штучної нейронної мережі з використанням власних даних – цей підхід полягає в тому, щоб взяти попередньо навчену модель і дотренувати її на власному наборі даних; це дозволяє швидше досягти високої точності, оскільки модель вже має загальне розуміння певних характеристик і може швидше адаптуватися до нових завдань [12].

Для тренування нейронної мережі на власних даних, необхідно мати анотації, тобто правильні мітки, які нейронна мережа має передбачити. Отримання таких анотацій зазвичай вимагає значних ресурсів. Модель, натренована на невеликому наборі даних, може страждати від перенавчання, що означає, що вона може показувати слабкі результати для примірників, які не входили в тренувальні дані, і має обмежену здатність до узагальнення на нові умови отримання зображення. Тому набір даних для тренування повинен бути різноманітним, включати різні варіації, такі як освітлення, кут повороту, розмір, колір, форма тощо. Самостійне тренування

моделі може вимагати значних обчислювальних ресурсів, проте це дозволяє обмежити класи лише до необхідних для конкретної задачі. Найоптимальнішим варіантом може бути дотренування вже готової моделі з використанням власних даних. Це дозволяє скористатися попередньо навченими вагами моделі та швидше досягти високої точності, адаптуючи модель до нових даних.

Глибинні нейронні мережі навчаються з нуля, використовуючи величезні набори даних, які містять мільйони зображень. Ці нейронні мережі зазвичай добре узагальнюються для великої кількості класів. Хоча використовуються лише декілька класів з усіх доступних, багато зображень об'єктів різних класів мають спільні риси. Функції одного детектора об'єктів одного класу можуть ефективно працювати під час пошуку об'єкта іншого класу. Це дає змогу перетренувати кілька останніх шарів нейронної мережі, вважаючи решту як вже побудований екстрактор ознак [13].

1.2 Методи глибокого навчання для виявлення ознак у послідовних даних

Для автоматизації процесу виявлення насильства під час відеоспостереження найбільш використовуваними моделями нейромереж є рекурентні нейромережі RNN (найбільш популярним видом даної категорії є Long Short-Term Memory (LSTM)) та згорткові нейронні мережі (CNN) [14].

Рекурентна нейронна мережа (RNN) є типом штучної нейронної мережі, що використовується для аналізу послідовних даних. Вона має здатність запам'ятовувати контекст з попередніх кроків і використовувати цю інформацію для обробки наступних кроків у послідовності. Основна особливість RNN полягає у зворотному зв'язку, що дозволяє передавати інформацію з попередніх кроків наступним. Це робить RNN особливо ефективним для аналізу і моделювання послідовних даних, таких як мовний текст, часові ряди або музичні сигнали, а також розпізнавання залежностей та шаблонів у цих даних. Структура RNN складається з повторюваних блоків, які називаються "рекурентними клітинками". Кожна клітинка приймає на вхід вхідний сигнал для поточного кроку і прихований стан з попереднього кроку. Після

обробки цих вхідних даних, вона видає новий прихований стан і вихідний сигнал, які передаються наступній клітинці в послідовності. Перевагою RNN є його здатність до моделювання довготривалих залежностей в послідовних даних. Він може враховувати контекст і залежності на різних рівнях в послідовності, що дозволяє йому вирішувати завдання, які потребують розуміння контексту [15]. На рисунку 1.1 зображено загальний вигляд рекурентної нейронної мережі (RNN).

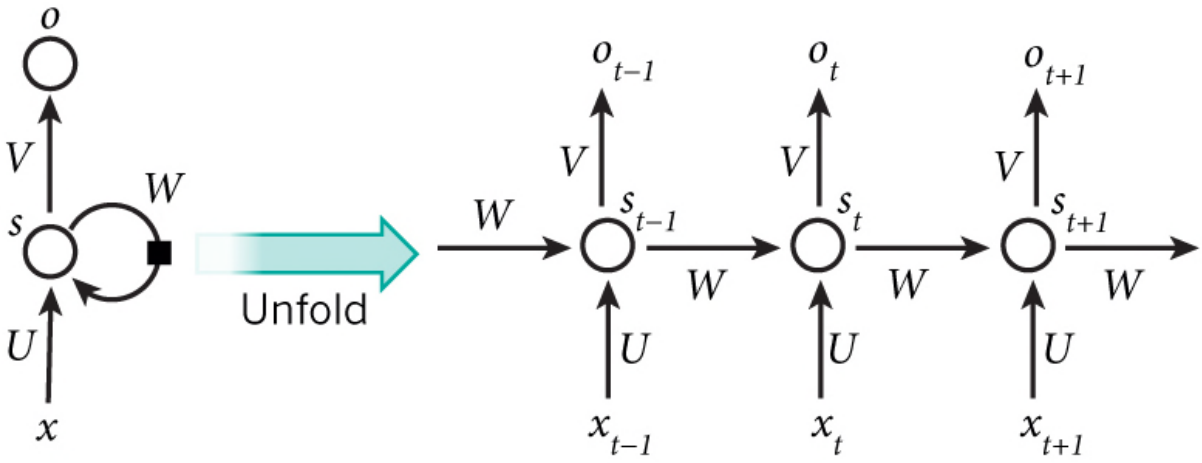


Рисунок 1.1 – Загальний вигляд рекурентної нейронної мережі [15]

CNN (згорткова нейронна мережа) є штучною нейронною мережею, яка використовується для аналізу візуальних даних, зокрема зображень. Вона є одним з найпоширеніших типів нейронних мереж, особливо в області комп'ютерного зору та обробки зображень. Основною особливістю CNN є використання згорток, які дозволяють автоматично виявляти локальні шаблони або ознаки у зображенні. Згортки є фільтрами, які проходять по зображенню, виконуючи операцію згортки, що полягає у перемноженні значень пікселів у вікні з відповідними вагами і підсумовуванні результатів. Цей процес дозволяє виділити різні ознаки, такі як гострі кути, краї, текстури тощо [16].

CNN складається з кількох шарів, включаючи згорткові шари, агрегаційні шари та повнозв'язані шари. Згорткові шари виявляють ознаки на різних рівнях абстракції, агрегаційні шари зменшують розмір зображення та кількість параметрів, а повнозв'язані шари виконують остаточний аналіз та класифікацію. CNN широко

використовується для багатьох завдань обробки зображень, таких як класифікація зображень, виявлення об'єктів, сегментація зображень, розпізнавання обличь, стилізація зображень та багато іншого. Вона також може бути використана в комбінації з іншими типами нейронних мереж для вирішення складніших задач аналізу даних. CNN є потужним і ефективним інструментом для обробки великих обсягів візуальних даних та визначення складних залежностей у зображеннях. Вона дозволяє автоматично виявляти ознаки, які людському оку можуть бути непомітними, і забезпечує високу точність у багатьох задачах комп'ютерного зору [17]. На рисунку 1.2 зображено загальний вигляд згорткової нейронної мережі (CNN).

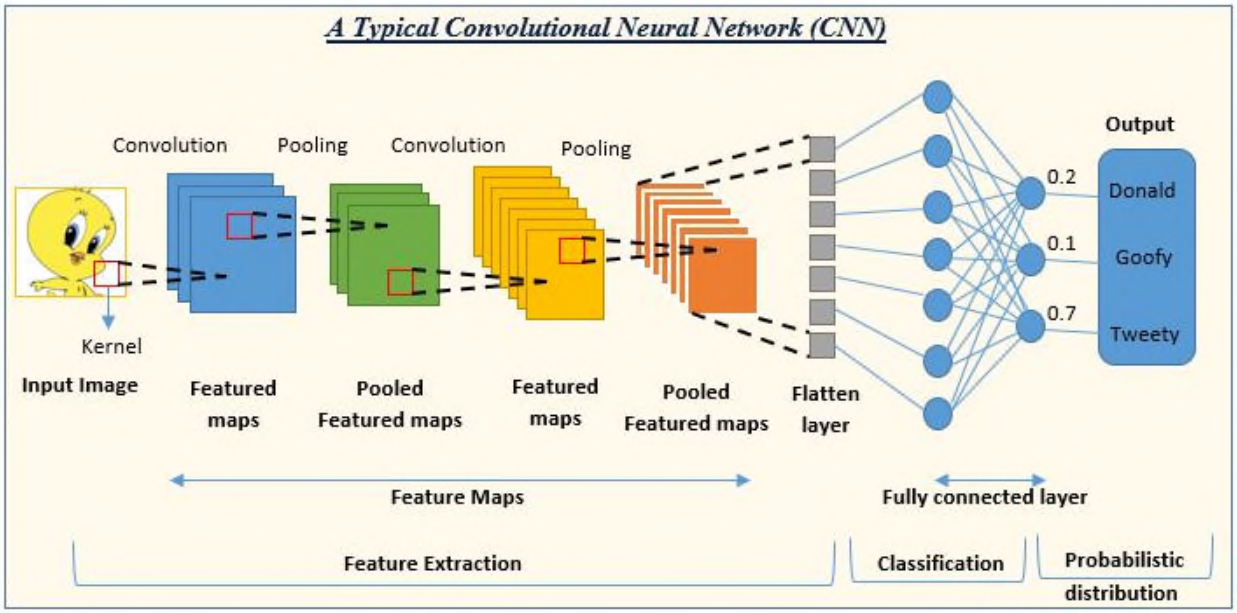


Рисунок 1.2 – Загальний вигляд згорткової нейронної мережі [17]

Дані моделі нейромереж рідко використовуються у “чистому” вигляді, на сьогодні їх розробляють у вигляді кастомних рішень, тобто розширяють та модифікують існуючу архітектуру або комбінують декілька архітектур. Прикладами таких нейромереж є Bidirectional Convolutional LSTM, 3D Convolutional Neural Network [18].

1.3 Аналіз існуючих публікацій виявлення проявів насильства нейромережевими засобами

Bidirectional Convolutional LSTM є комбінацією двох популярних архітектур нейронних мереж: Convolutional Neural Network і LSTM, які використовуються для обробки послідовностей, зокрема зображень і відео. Ця архітектура поєднує в собі здатність CNN виділяти важливі ознаки зображення та здійснювати конволюційні операції зі зміщеннями (передбачаючи просторову структуру), і здатність LSTM взаємодіяти з попередніми даними в послідовності для управління пам'яттю та довгостроковими залежностями. Bidirectional Convolutional LSTM широко використовується в завданнях обробки відео, де важливий контекст як в просторовому, так і в часовому вимірах.

Основна ідея полягає в тому, щоб використовувати CNN для виділення ознак з кадрів відео та застосовувати Bidirectional LSTM для аналізу послідовності цих ознак в часі. Це дозволяє моделі розуміти довгострокові залежності в відеоданих і виявляти складні структури, такі як рух об'єктів чи дії відносно часу. Bidirectional Convolutional LSTM знайшла застосування в таких завданнях, як розпізнавання дій відносно часу, відслідковування об'єктів в відео, аналіз медичних зображень та багатьох інших в сферах, де важливий контекст як у просторовому, так і в часовому вимірах [19].

Використання моделі Bidirectional Convolutional LSTM наведено у роботі «Bidirectional Convolutional LSTM for the Detection of Violence in Videos» [20], де для вирішення задачі по виявленню насильства на відео був запропонований підхід, який полягає у створенні методу в якому поєднанні декілька архітектур нейромереж: Bidirectional Convolutional LSTM та VGG13. На рисунку 1.3 зображено архітектуру даного методу.

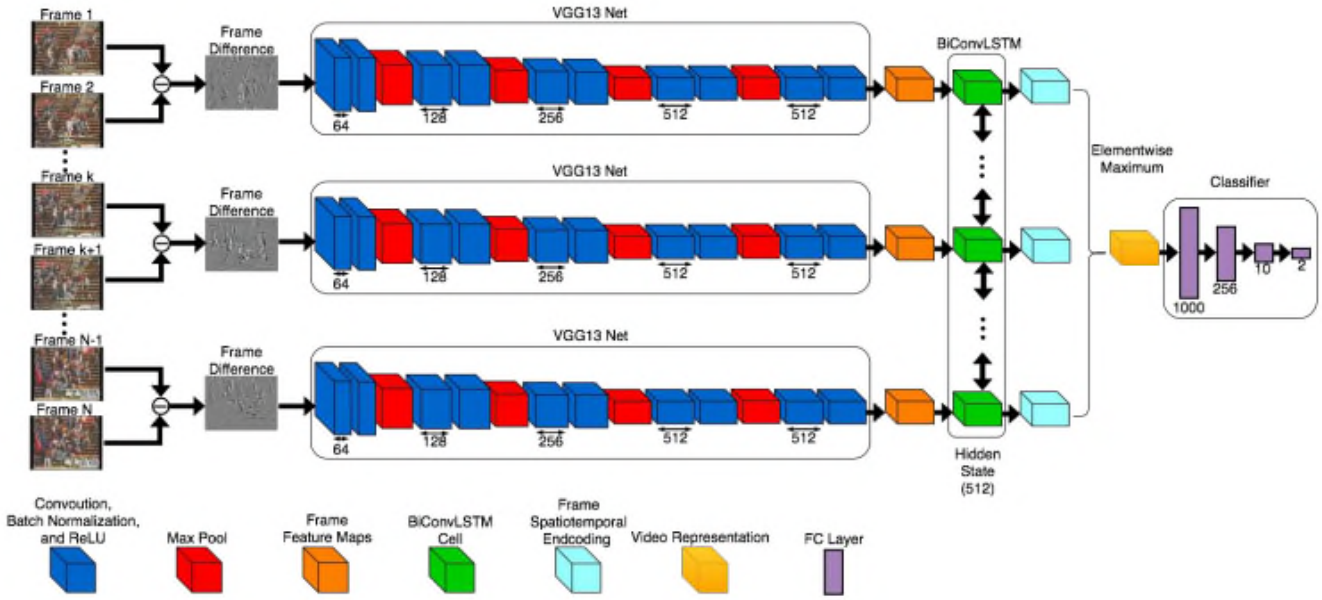


Рисунок 1.3 – Архітектура методу «Просторово-часовий кодер» [20]

Принцип роботи запропонованого методу можна розділити на 3 етапи:

Етап 1 – Фільтрація відео

Перший етап передбачає розділення відео на окремі фрейми та їх подальше фільтрування. Використовуються різні фільтри для видалення шуму та виділення важливих візуальних ознак. Застосування фільтрації допомагає покращити якість вхідних даних та підготувати їх для подальшого аналізу. Для фільтрації вхідного відеоматеріалу використовуються нейромережа VGG13.

Етап 2 – Відбір ознак

Наступний етап полягає у виборі ознак для аналізу, які найкраще корелюють з наявністю насильства в відео. Ознаки можуть включати кольорові характеристики, текстурні особливості та інші візуальні параметри. Вибір правильних ознак важливий для точності класифікації. Для виділення ознак із відфільтрованих даних використовується модель Bidirectional Convolutional LSTM.

Етап 3 – Класифікація

Оброблені ознаки подаються на бінарну рекурентну нейронну мережу, яка визначає, чи є відео насильством. Бінарна рекурентна нейронна мережа складається з двох рекурентні вузлів, які обробляють відео в прямому і зворотному напрямках.

Кожен вузол складається з послідовності згорткових шарів, які обробляють послідовність фреймів. На виході кожного вузла отримується вектор ознак, який потім обробляється лінійним шаром для отримання прогнозу.

Тестування запропонованого методу проводилося на 3 наборах даних: «Hockey Fights», «Violent Flows», «Movies». Загальна точність тесту для набору даних «Hockey Fights» сягає 96,96 ± 1.08%, для набору даних «Violent Flows» сягає 92,18 ± 3,29%, для «Movies» сягає 100% [20].

Іншим підходом для вирішення задачі по виявленню насильства на відео є використання модифікованої версії згорткової нейронної мережі. 3D CNN (3D Convolutional Neural Network) представляє собою варіант глибокої нейронної мережі, яка використовується для обробки відео та тривимірних даних. Вона є розширенням звичайних 2D CNN, які використовуються для обробки зображень. 3D CNN здатна аналізувати просторові залежності та динаміку у тривимірних даних, таких як відео, медичні зображення, сенсорні дані тощо. Основна ідея 3D CNN полягає в застосуванні тривимірних згорткових шарів до вхідних даних. Кожен згортковий шар складається з набору фільтрів, які здійснюють згортку із вхідних даних. Згортка полягає в переміщенні фільтра по вхідних даних та обчисленні скалярного добутку між фільтром та відрізком даних, що відповідає розміру фільтра. Цей процес дозволяє виявляти локальні особливості та шаблони в тривимірних даних. 3D CNN може бути використана для різних завдань, таких як класифікація відео, детекція об'єктів, розпізнавання дій та багато іншого. Вона є потужним інструментом для розуміння тривимірних даних та виконання складних завдань у сфері комп'ютерного зору та обробки відео [21].

Використання даного підходу наведено у роботі «Efficient Violence Detection Using 3D Convolutional Neural Networks» [22]. Реалізація запропонованого методу зображена на рисунку 1.4.

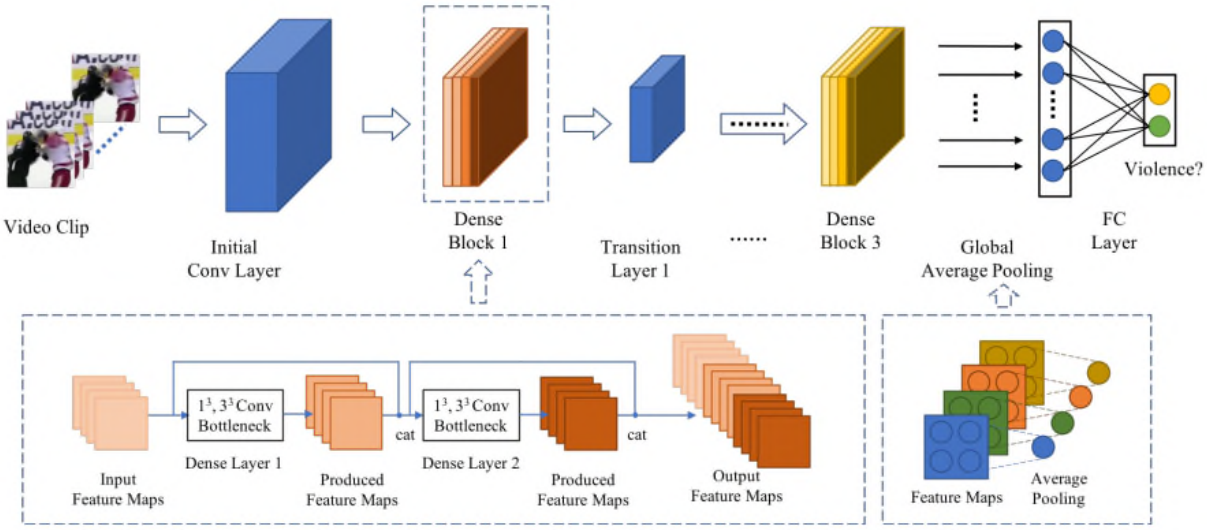


Рисунок 1.4 – Реалізація методу виявлення насильства за допомогою 3D CNN [22]

Роботу даного методу можна розділити на 3 етапи:

Етап 1 – Фільтрація даних

Початковий етап передбачає розділення відео на окремі фрейми та їх подальше фільтрування. На даному етапі дані про відео фільтруються для видалення шумів і нерелевантної інформації. Для фільтрації даних використовуються різні фільтри, такі як фільтр низьких частот, фільтр високих частот і фільтр Гауса.

Етап 2 – Виділення просторово-часових ознак

На другому етапі виділяються просторово-часові ознаки з відео. Просторово-часові ознаки являються характеристиками, які описують зміну яскравості в часі та просторі. Для виділення просторово-часових ознак використовуються 3D-конволюційна нейронна мережа. 3D-конволюційна нейронна мережа складається з шарів, які обробляють дані в тривимірному просторі.

Етап 3 – Класифікація відео

На третьому етапі відео класифікуються як насильство або не насилля на основі виділених просторово-часових ознак. Класифікація здійснюється шляхом застосування логістичної регресії до вилучених ознак, отриманих на другому етапі.

Для тестування запропонованого методу було використано три загальнодоступних наборів даних, таких як «Violent Flows», «Hockey Fights» та

«Movies». Точність для набору даних «Violent Flows» сягає 97.17±0.95%, для «Hockey Fights» сягає 98.3±0.81%, для «Movies» сягає 100±0% [22].

Отже, проаналізувавши наведенні наукові публікації можна зробити наступні висновки:

Bidirectional Convolutional LSTM є потужною моделлю, який включає в себе комбінацію рекурентних LSTM шарів та конволюційних шарів. Він дозволяє ефективно розпізнавати часові залежності в послідовностях даних та має високу стійкість до змін у часовому контексті. Дана модель особливо корисна для завдань, де важливі динамічні зміни в даних, такі як розпізнавання жестів або відео аналітика.

З іншого боку, 3D CNN є моделлю, спеціально призначена для аналізу тривимірних даних, таких як відео. Вона може здійснювати згортку не тільки по просторових, але й по часових вимірах, враховуючи динаміку подій в часі. Це робить 3D CNN дуже ефективним для завдань, де важлива взаємодія об'єктів в тривимірному просторі, наприклад, в розпізнаванні дій у відеозаписах.

У наведених роботах для вилучення ознак із вхідного відеоматеріалу використовується модель згорткової нейронної мережі Модель CNN є необхідною в розпізнаванні через її здатність виявляти локальні особливості та шаблони в даних, безпосередньо взаємодіючи з ними. Завдання розпізнавання, такі як класифікація об'єктів або дій, зазвичай потребують виявлення локальних особливостей, таких як краї, текстури або форми. CNN може автоматично вивчати ці особливості шляхом застосування фільтрів на різних рівнях абстракції, що дозволяє виявляти більш складні залежності між пікселями або вузлами.

Тому використання архітектури нейромережі CNN у розробці кастомного рішення є цілком виправдним та необхідним для того, щоб реалізовувати ефективне рішення для розв’язку поставленого завдання.

1.4 Постановка задачі

Метою кваліфікаційної роботи магістра є розробка методу виявлення зовнішніх проявів насильства у відеопотоці нейромережевими засобами. Тому для виконання поставленого завдання потрібно:

- Провести аналіз нейромережевих моделей та існуючих підходів для виявлення проявів насильства у відеопотоці;
- Розробити метод виявлення зовнішніх проявів насильства у відеопотоці з використанням згорткової нейронної мережі та класифікатора SVM;
- Підготувати набір даних для навчання згорткової нейронної мережі;
- Навчити попередньо навчену згорткову нейронну мережу виявляти ознаки насильства на неперервному відеопотоці даних;
- Визначити загальну точність запропонованого методу виявлення зовнішніх проявів насильства.

Вдалим виконанням завдання можна вважати високу точність методу у процесі виявлення насильства.

Висновки до розділу 1

Після опрацювання першого розділу було сформовано мету роботи, завдання дослідження та доведено актуальність роботи. Проведено аналіз та детально описано предметну область, наведено 2 публікації, які пов'язані з даною темою, що свідчить про успішну апробацію запропонованого методу вирішення задачі. Описано постановку завдання, в якій, по пунктах, наведено задачі, які необхідно вирішити в даній роботі.

РОЗДІЛ 2

Розробка методу виявлення зовнішніх проявів насильства у відеопотоці нейромережевими засобами

Сьогодні реалізація ефективного рішення для розв’язування поставленого завдання базується на модифікації існуючих рішень або поєднанні декількох рішень з метою отримання ефективного засобу розв’язку поставленої задачі. Для реалізації методу виявлення зовнішніх проявів насильства у відеопотоці нейромережевими засобами за основу взято математичну модель згорткової нейронної мережі [23].

Реалізація методу виявлення зовнішніх проявів насильства у відеопотоці нейромережевими засобами складається з 6 етапів:

– *Етап 1 – Підготовка вхідних даних.* Реалізація механізму вилучення із вхідного відеоматеріалу набору кадрів та налаштування відповідних параметрів для кадрів;

– *Етап 2 – Реалізація шарів згортки.* Виконання операції згортки над вхідними даними з метою формування карт ознак;

– *Етап 3 – Реалізація агрегаційних шарів.* Виконання операції максимального об’єднання з метою зменшення розмірності даних та одночасним збереженням найважливіших характеристик;

– *Етап 4 – Реалізація повнозв’язного шару.* Виконання процесу преобразування вихідних даних з попередніх шарів у векторну форму;

– *Етап 5 – Налаштування класифікатору SVM*. Формування оцінки, яка представляє ймовірність того, що вхідні дані належать до певного класу (насильницького або не насильницького) за допомогою гіперплощини в векторному просторі;

– *Етап 6 – Результат*. Відображення на відеоматеріалі у верхньому лівому кутку прямокутника відповідного кольору та оцінки (якщо оцінка менше рівна 0.5 то зелений колір (ненасильство), в іншому випадку червоний (насильство)).

На рисунку 2.1 зображено схему методу виявлення зовнішніх проявів насильства у відеопотоці нейромережевими засобами.

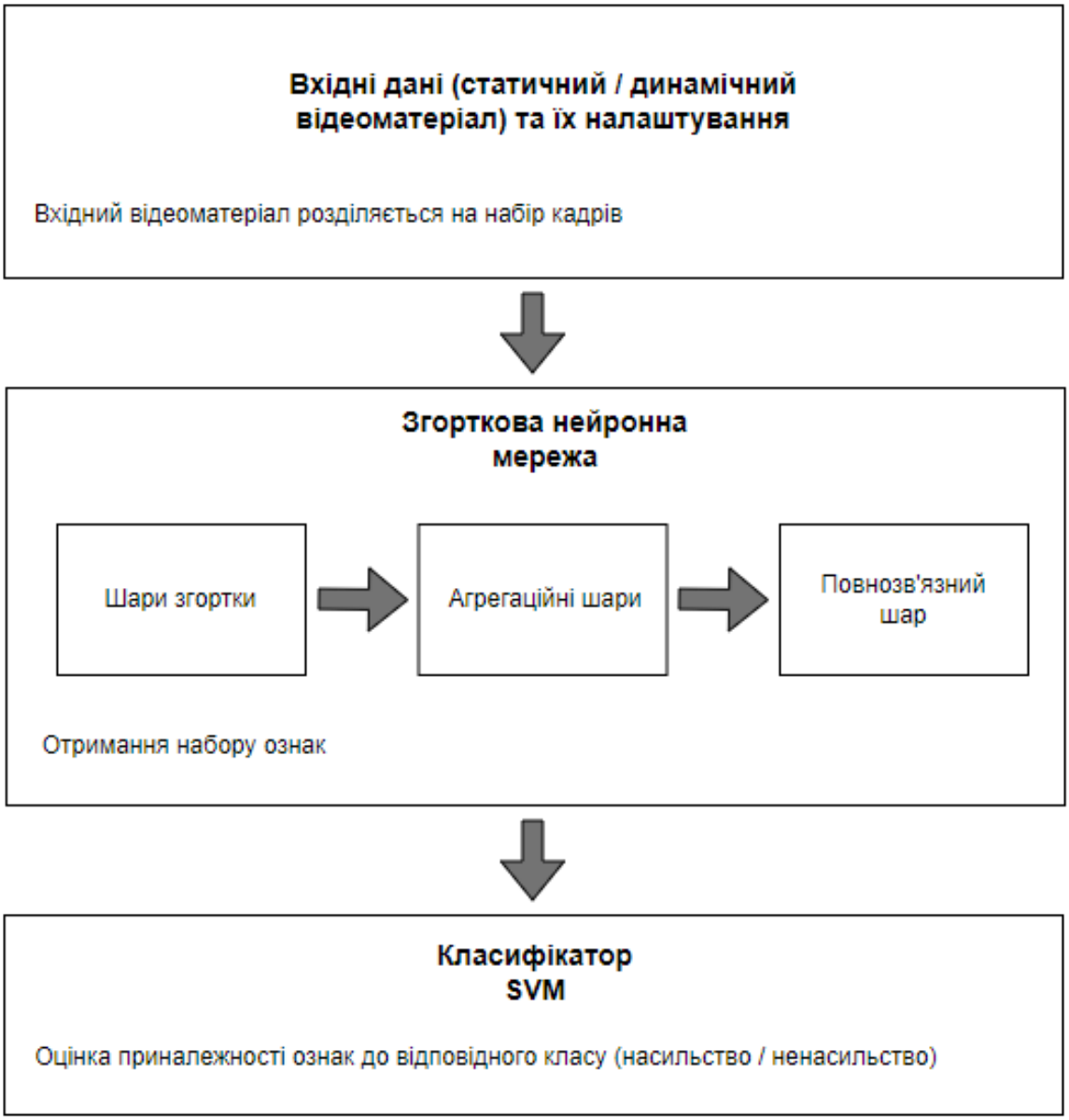


Рисунок 2.1 – Схема методу виявлення зовнішніх проявів насильства у відеопотоці нейромережевими засобами

2.1 Згорткова нейронна мережа запропонованого методу

На рисунку 2.2 зображено архітектуру методу для вирішення задачі виявлення зовнішніх проявів насильства у відеопотоці.

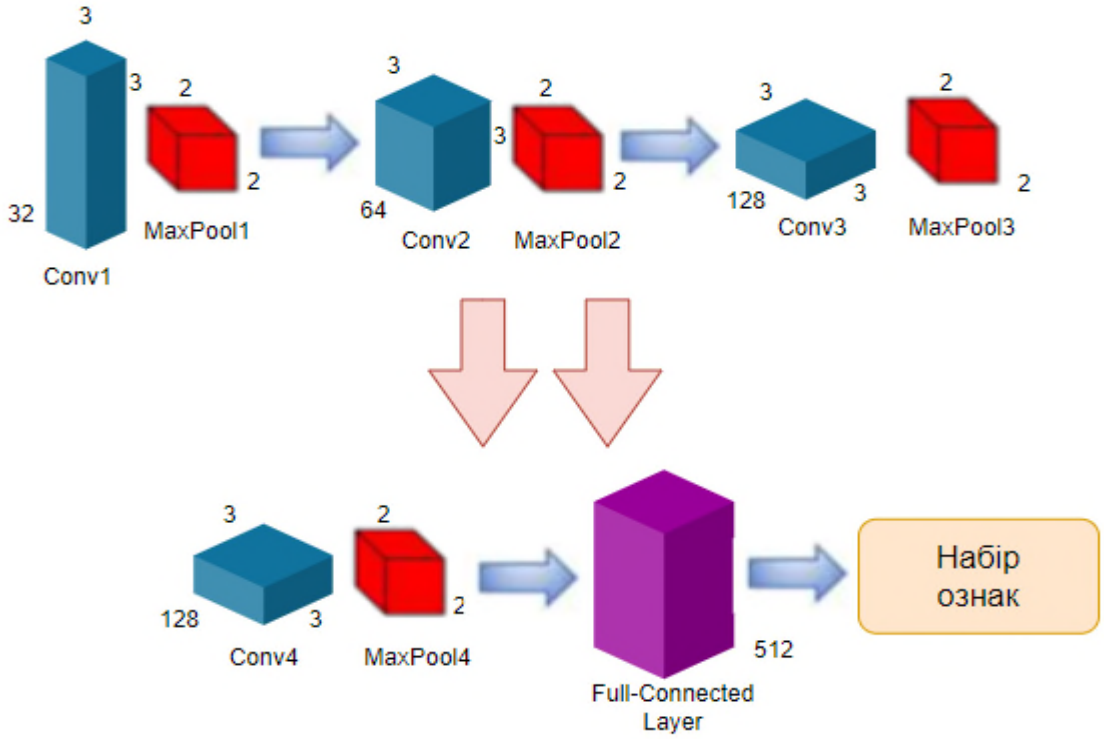


Рисунок 2.2 – Архітектура згорткової нейронної мережі в контексті запропонованого методу

Архітектура згорткової нейронної мережі складається з наступних шарів:

– Conv1. Шар згортки, який складається з 32 фільтрів розмірністю 3x3;
– MaxPool1. Шар максимального об’єднання з розмірністю фільтра 2x2;
– Conv2. Шар згортки, який складається з 64 фільтрів розмірністю 3x3;
– MaxPool2. Шар максимального об’єднання з розмірністю фільтра 2x2;
– Conv3. Шар згортки, який складається з 128 фільтрів розмірністю 3x3;
– MaxPool3. Шар максимального об’єднання з розмірністю фільтра 2x2;
– Conv4. Шар згортки, який складається з 128 фільтрів розмірністю 3x3;
– MaxPool4. Шар максимального об’єднання з розмірністю фільтра 2x2;
– Full-Connected layer. Повнозв’язний шар, який складається з 512 нейронів.

Кінцевим результатом згорткова нейронна мережа видає набір ознак у вигляді набору векторів, який потім передається на вхід класифікатору SVM.

2.1.1 Налаштування вхідних даних

В якості вхідних даних у згорткову нейромережу передається відеоматеріал. Відео розбивається на послідовність пакетів по 20 кадрів, де кожен кадер має розмір 150x150 пікселів. Розбиття відео на послідовність пакетів кадрів та встановлення відповідних параметрів до кадрів дозволяє нейромережі ефективніше з точки зору пам’яті та швидше працювати з даними. Після формування пакетів кадрів нейромережа починає опрацьовувати кожен кадр окремо з метою визначення ознак, які дозволять визначити чи несе відео насильницький характер.

2.1.2 Шари згортки

Першим кроком в отриманні ознак виступає механізм згортки. Згортка є математичною операцією, яка використовуючи ядро згортки (детектор ознак) на вхідне зображення, в якості результату формує карту ознак.

Детектор ознак зазвичай складається з певної фіксованої кількості фільтрів, кожен з яких, з математичної точки зору, є матрицею параметрів. Розмір цих фільтрів зазвичай обмежений таким чином, щоб вони покривали лише невелику частину зображення по ширині і висоті, але мали ту саму розмірність у глибину, що й вхідне зображення. Для даної архітектури було обрано розмірність фільтрів 3x3, оскільки більші розмірності призводять до збільшення обчислювальних витрат, тоді як менші розмірності дають дрібнозернисті та локальні ознаки, пропускаючи інформацію з сусідніх пікселів. Одним з важливих факторів вибору розмірності фільтрів є їх парність або непарність. Фільтри непарного розміру симетрично розділяють пікселі попереднього шару навколо вихідного пікселя. Якщо ця симетрія порушена, можуть виникнути спотворення між шарами, тому фільтри розміром 2x2 або 4x4 не є оптимальними. На рисунку 2.3 зображено детектор ознак.

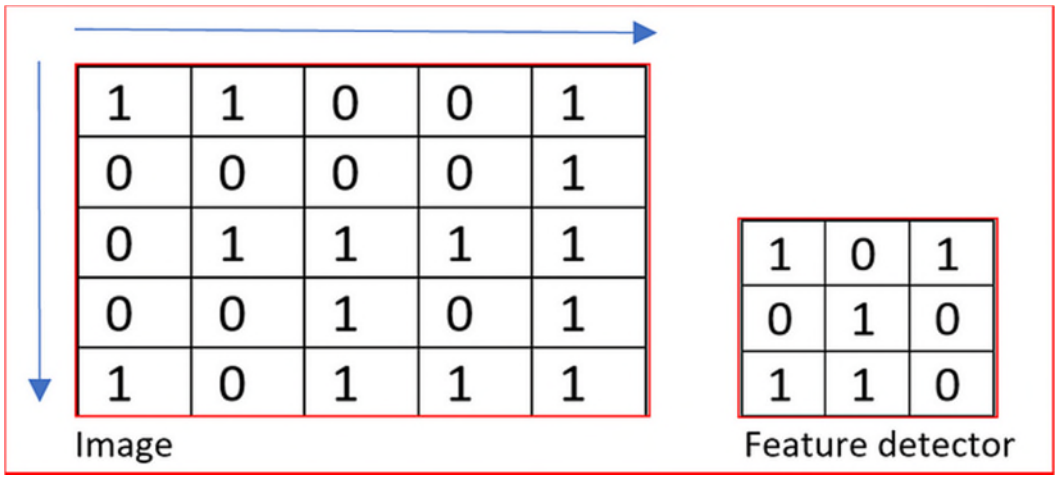


Рисунок 2.3 – Детектор ознак [23]

Кількість фільтрів являється одним із гіперпараметрів нейронної мережі, оскільки збільшення кількості фільтрів веде до збільшення кількості потенційно виявлених ознак та закономірностей, але при цьому занадто велика їх кількість може призвести до повторень та знаходження ознак, які не є характеристичними, тобто до ускладнення та потенційного перенавчання моделі. Тобто, кількість фільтрів відповідає кількості карт ознак, яка визначає глибину вихідних даних, а, отже, і глибину вхідних для наступного рівня мережі. В архітектурі даної нейромережі представлено наступну кількість фільтрів: 32, 64, 128, 128.

Визначення коректної кількості фільтрів залежить від задачі яку потрібно вирішити і визначити коректну кількість можна лише опираючись на результати навчання нейромережі. Для вирішення задачі виявлення зовнішніх проявів насильства у відеопотоці обрано велику кількість фільтрів, яка на кожному рівні нейромережі збільшується удвічі. Велика кількість фільтрів дозволяє отримати більше ознак, але надмірна кількість може призвести до отримання лишніх не характерних ознак, а також до перенавчання моделі. Збільшення фільтрів на кожному рівні має свою особливість. Кожен шар фільтрів призначений для захоплення шаблонів, наприклад, перший шар фільтрів фіксує для прикладу, краї, кути, крапки тощо. Наступний рівень шарів об’єднують ці візерунки, щоб створити більші візерунки (наприклад, квадрати, кола тощо). Тобто на кожному рівні формуються все більш складні об’єкти, що в кінцевому результаті призведе до формування

повноцінного образу. Ось чому збільшується кількість фільтрів в наступних шарах, щоб охопити якомога більше комбінацій.

Крок обходу зображення є одним із гіперпараметрів нейронної мережі, тобто параметром, що має бути визначений ще на етапі конструювання моделі. Збільшення кроку веде до зменшення розмірності вихідних даних (а, отже, до зменшення складності моделі). Для даної мережі обрано крок обходу 1, так як завелике значення, в даному випадку, може призвести до втрати значних ознак, що в свою чергу буде негативно впливати на кінцевий результат.

Доповнення нулями також є одним із гіперпараметрів мережі і може використовуватися в тих випадках коли небажано зменшувати розмірність карт ознак у порівнянні з вхідними даними. Суть даного методу полягає у розширенні вхідного зображення методом додавання певної кількості нульових значень з обох сторін. Це дозволяє збільшити вплив значень, що розташовані по краях зображення. Без використання цього методу, крайні значення використовуються лише один раз у карті ознак, що може призвести до втрати інформації з країв зображення. На рисунку 2.4 зображено операцію доповнення нулями.

0	0	0	0	0	0	0
0	1	1	0	0	1	0
0	0	0	0	0	1	0
0	0	1	1	1	1	0
0	0	0	1	0	1	0
0	1	0	1	1	1	0
0	0	0	0	0	0	0

Рисунок 2.4 – Операція доповнення нулями [23]

Результатом виконання операції згортки є формування карти ознак. Вона представляє собою просторове відображення важливих ознак, що були виявлені мережею під час процесу навчання. Основна ідея полягає в тому, що кожен шар згортки виконує операцію згортки над вхідними даними, використовуючи набір фільтрів, що виявляють різні ознаки. Кожне ядро здійснює згортку з вхідним зображенням та створює карту ознак, яка підкреслює присутність цих ознак у вихідних даних. На кожний фільтр формується своя карта ознак. Карти ознак отримані на попередньому шарі згортки зазвичай використовуються в якості вхідних даних для наступних шарів згортки, дозволяючи знаходити більш складні ознаки. Для знаходження карт ознак використовуються математична формула (1), на рисунку 2.4 зображено виконання даної операції:

$$\mathrm{M(i,j)} = (\mathrm{K} * \mathrm{X})(\mathrm{i,j}) = \sum_{m}\sum_{n} \mathrm{K(m,n)\,X(i-m,j-n)}, \quad (1)$$

де M – елемент карти ознак з координатами i та j,

X – вхідне зображення,

K – детектор ознак,

m, n – розмірності детектора ознак.

Розмірність карт ознак залежить від вказаних гіперпараметрів та може бути обчислена за допомогою формули (2):

$$d_m = \frac{d_x - f + 2 * p}{s + 1}, \quad (2)$$

де d_m – розмірність карт ознак (ширина / висота),

d_x – відповідна розмірність вхідних даних,

f – розмірність фільтрів,

p – доповнення нулями,

s – розмір кроку обходу зображення.

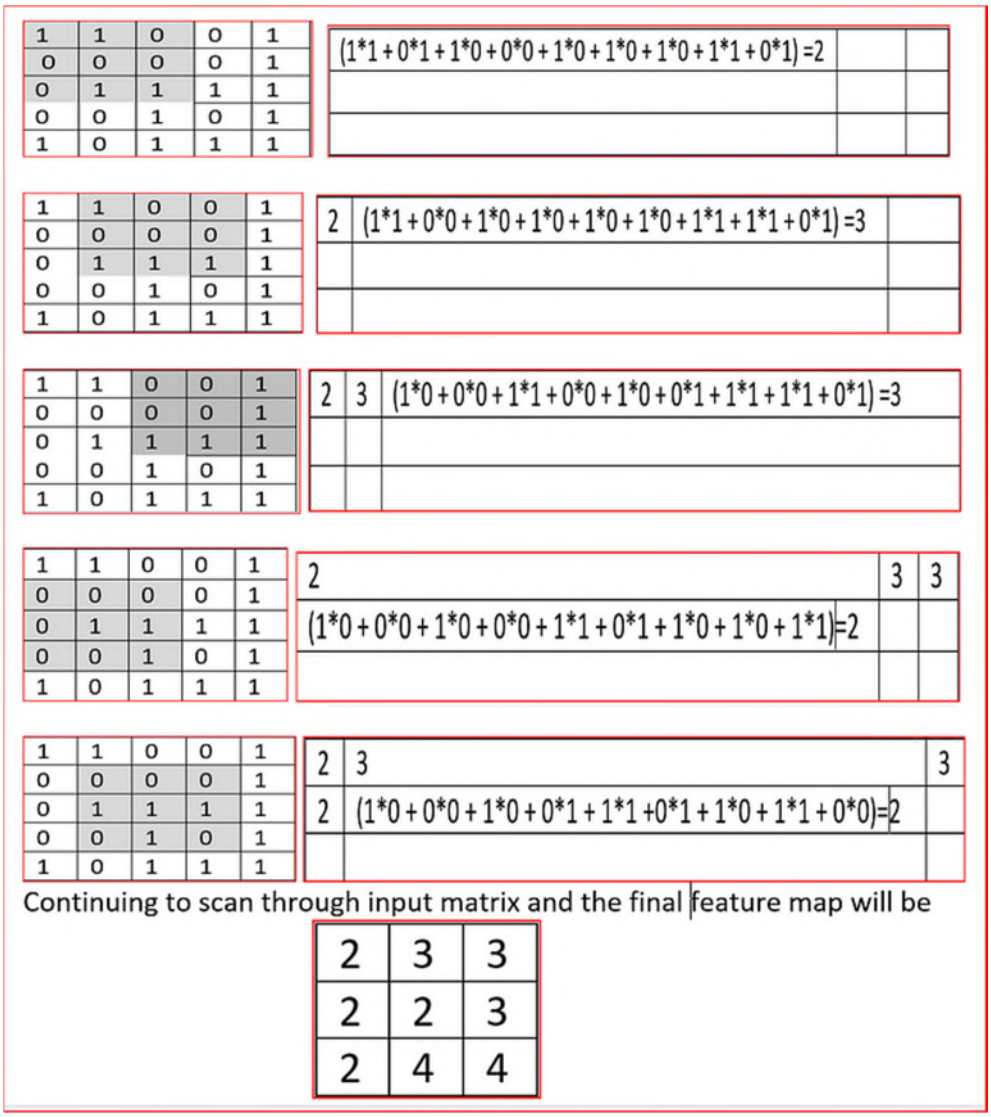


Рисунок 2.5 – Формування карти ознак [23]

Після формування карт ознак до отриманих даних необхідно застосувати функцію активації. Функція активації є функцією, яка перетворює вхідні лінійні дані у вихідні не лінійні. Вибір та використання функції активації має значний вплив на функціонування мережі. Без використання цієї функції мережа не змогла б навчатися розпізнавати нелінійні закономірності та розподіли.

В якості функції активації для згорткового шару було використано функцію ReLU (випрямлена лінійна одиниця). ReLU є активаційною функцією, яка оброблює дані та навчає мережу значно швидше, ніж інші активаційні функції, одночасно

додаючи не лінійність. Окрім прискорення навчання мережі, вона також зменшує ризик перенавчання, оскільки мережа навчається більш складним зв'язкам за той же час, порівняно з іншими активаційними функціями, що в свою чергу дозволяє навчати складніші моделі на одних і тих самих даних. За допомогою формули (3) можна знайти значення ReLU:

$$y = f(z) = \max(0, z), \tag{3}$$

де y – елемент поточного рівня,

z – елемент з вхідних даних.

Формула функції вказує на те, що вона усуває від'ємні значення у мережі, призводячи їх до нульового значення, та залишає незмінними інші значення, проводячи тотожне перетворення. Ефективність цієї функції з апаратної точки зору підвищується через її здатність просто порівнювати значення з нулем. Похідна функції ReLU приймає значення 0 для від'ємних вхідних значень і 1 для всіх інших, що спрощує процес зворотного поширення помилки в алгоритмі. На рисунку 2.6 зображено функцію активації ReLU.

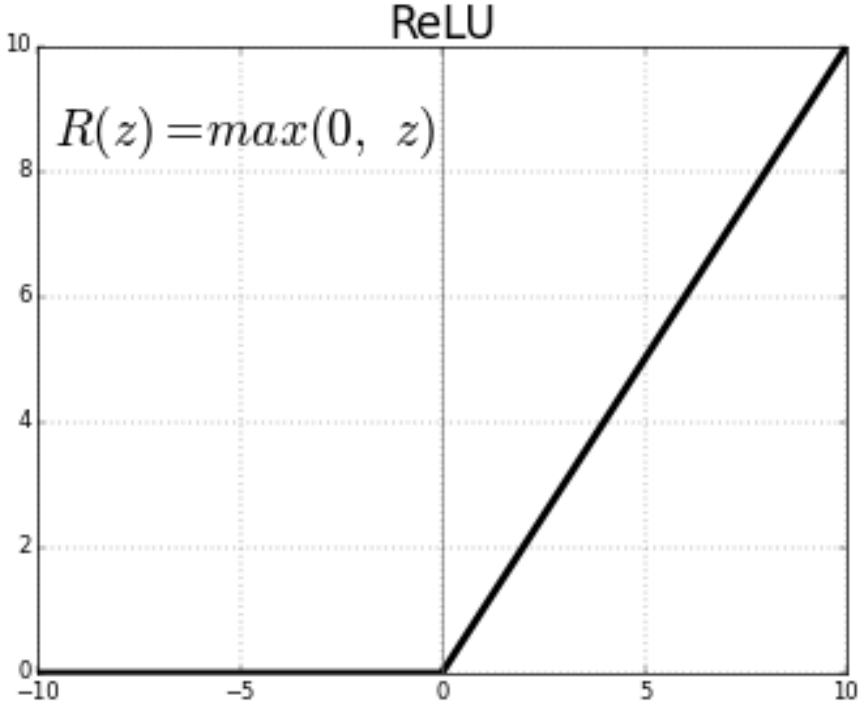


Рисунок 2.6 – Функція активації ReLU [23]

2.1.3 Агрегаційні шари

Наступним кроком у виявленні ознак є робота агрегувальних шарів. Агрегувальні шари є важливою частиною згорткових нейронних мереж, нарівні з згортковими шарами. Даний тип шарів в якості вхідних даних приймає результати виконання згортки. Основне завдання агрегаційних шарів полягає у зменшенні розмірності даних, з одночасним збереженням найважливіших характеристик, шляхом формування залежності між кількома елементами з попереднього шару, з єдиним елементом поточного шару. Тому при побудові мережі зазвичай використовують агрегаційні шари з певною періодичністю між згортковими шарами.

Агрегувальні шари в комп'ютерних нейронних мережах можуть бути розділені на два підтипи: усереднювальні та максимізаційні. Усереднювальні шари використовуються для обчислення середнього значення серед елементів, що належать до відповідних груп нейронів з попереднього шару, тоді як максимізаційні шари надають максимальне значення. Важлива роль агрегувальних шарів полягає в забезпеченні стійкості мережі до змін вхідних даних, таких як зміна кута погляду на об'єкт або його положення. Це досягається шляхом вибору значення з множини відповідно до певного критерію, яким може бути усереднювання або максимізація. Навіть незначні зміни вхідних даних майже не впливають на результати агрегаційного шару, що дозволяє отримувати подібні вихідні значення для схожих ознак, таких як переходи, кути або заокруглення, незалежно від змін в положенні або освітленні.

Для вирішення поставленої задачі в якості шару агрегування було обрано максимізацію, так як даний тип себе краще демонструє у роботі з відповідними ознаками: кути, заокруглення, що в контексті поставленої задачі має важливе значення. Максимізація (максимальне об’єднання) представляє собою операцію об'єднання, під час якої вибирається найбільший елемент з області, охопленої фільтром, на карті об'єктів. Це означає, що після проходження через шар максимального об'єднання отримується карта ознак, яка містить найвидатніші ознаки

з попередньої карти ознак. Таким чином, максимізація дозволяє виділити найбільш важливі ознаки та особливості на карті об'єктів. На рисунку 2.7 зображено приклад роботи максимального об’єднання, також наведено формулу (4) для реалізації даної операції:

$$p(i,j) = max_{i,j}\big(x(i-m, j-n)\big), \tag{4}$$

де p(i, j) – значення елементу поточного рівня з координатами і та j,

x – вхідні дані з попередніх рівнів,

m, n – розмірність рецептивного поля.

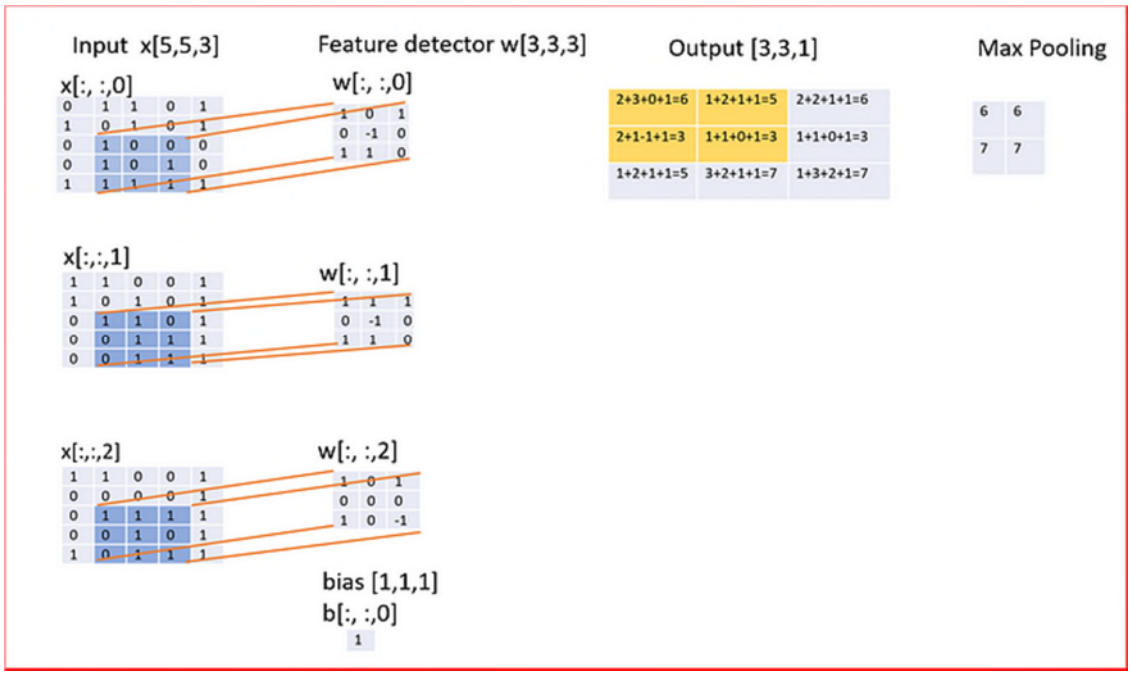


Рисунок 2.7 – Виконання операції максимального об’єднання [23]

Агрегаційні шари мають кілька однакових гіперпараметрів, які використовуються разом з операцією згортки. Ці параметри включають визначення кроку обходу зображення, який часто обирається таким чином, щоб уникнути накладання; доповнення нулями, а також розмір рецептивного поля. Розмірність результуючих даних для агрегаційних шарів може бути обчислена з використанням

значень гіперпараметрів, аналогічно обчисленню розмірності карт ознак для згорткових рівнів. Для представленої архітектури нейромережі було обрано розмірність рецептивного поля 2x2. Агрегаційні шари спроможні значно зменшити розмірність вхідних даних та вплив окремих елементів з невеликими значеннями на результат. Тому використання такого малого рецептивного поля є обґрунтованим.

2.1.4 Повнозв'язний шар

Наступним кроком у виявленні зовнішніх проявів насильства у відеопотоці є робота повнозв'язного рівня згорткової нейронної мережі. Повнозв'язний рівень фактично представляє собою модель багаторівневого персептрона. Повнозв'язні рівні в згорткових нейронних мережах відносяться до рівнів, де всі нейрони з наступного шару з'єднані з нейронами попереднього шару, аналогічно більшості шарів у звичайних нейронних мережах. Повнозв'язні рівні зазвичай використовуються на передостанньому етапі роботи мережі для підготовки результатів на виході мережі. Проте, використання повнозв'язних рівнів на початкових та прихованих рівнях мережі може бути не обґрунтованим, оскільки це ускладнює модель і навіть може пропустити раніше виявлені ознаки та закономірності за допомогою згорток та агрегації. Для роботи повнозв'язного рівня елементи матриць та шарів з попереднього кроку випрямляються в послідовність значень. Такі шари, подібно до згорткових, виконують обчислення скалярного добутку даних та параметрів з додаванням зсуву. Наведено формулу (5) для повнозв'язного рівня:

$$y = f\left(w_{ij} * x + b\right), \tag{5}$$

де w – масив ваг,

x – вхідні дані з попереднього рівня,

b – зміщення,

f – функція активації.

На повнозв’язних рівнях зазвичай застосовуються активаційні функції, результат яких може бути використаний на наступному рівні або обчислений для отримання результатів мережі на виході, на рисунку 2.8 зображено вигляд повнозв’язного рівня.

Тому типи активаційних функцій, які використовуються на цьому рівні, зазвичай відрізняються від тих, що використовуються на згорткових та агрегаційних рівнях, і можуть варіюватися залежно від положення рівня та загальної мети мережі. Для повнозв’язного рівня представленої нейронної мережі було обрано функцію активації сигмоїди. Функція активації сигмоїди представляє собою не лінійну функцію, яка стискає вхідні дані до діапазону від 0 до 1. Функція активації сигмоїди, як правило, використовується у задачах бінарної класифікації, де потрібно передбачити можливість приналежності до одного з двох класів. Дана функція конвертує вхідні значення у ймовірності, які можуть бути інтерпретовані як імовірність приналежності до позитивного класу. Наведено формулу (6) для обрахування функції активації сигмоїд, на рисунку 2.8 зображено графічний вигляд даної функції активації:

$$y = f(x) = \frac{1}{1 + e^{-x}}, \tag{6}$$

де x – вхідні значення.

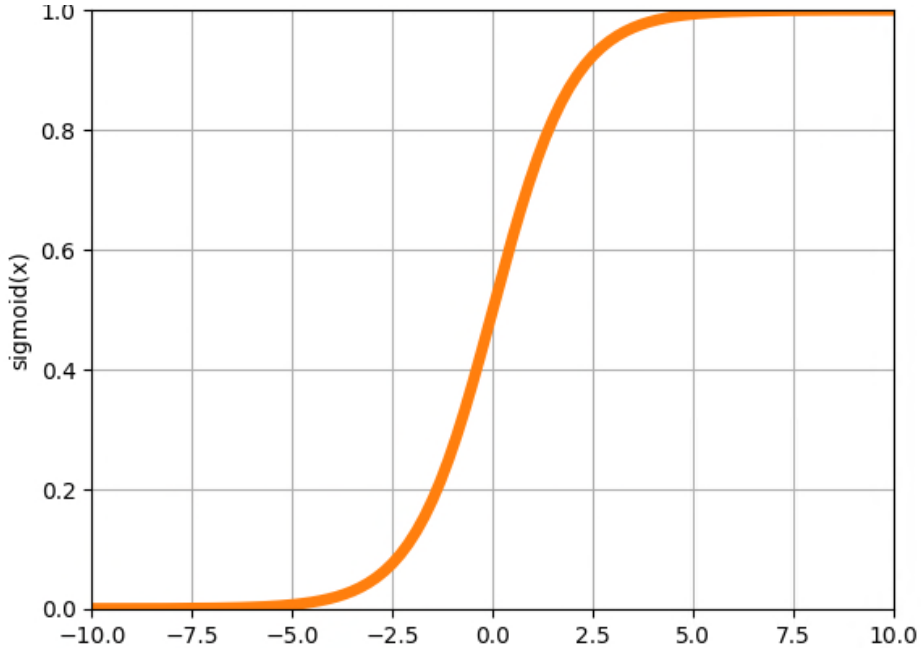


Рисунок 2.8 – Функція активації сигмоїд [23]

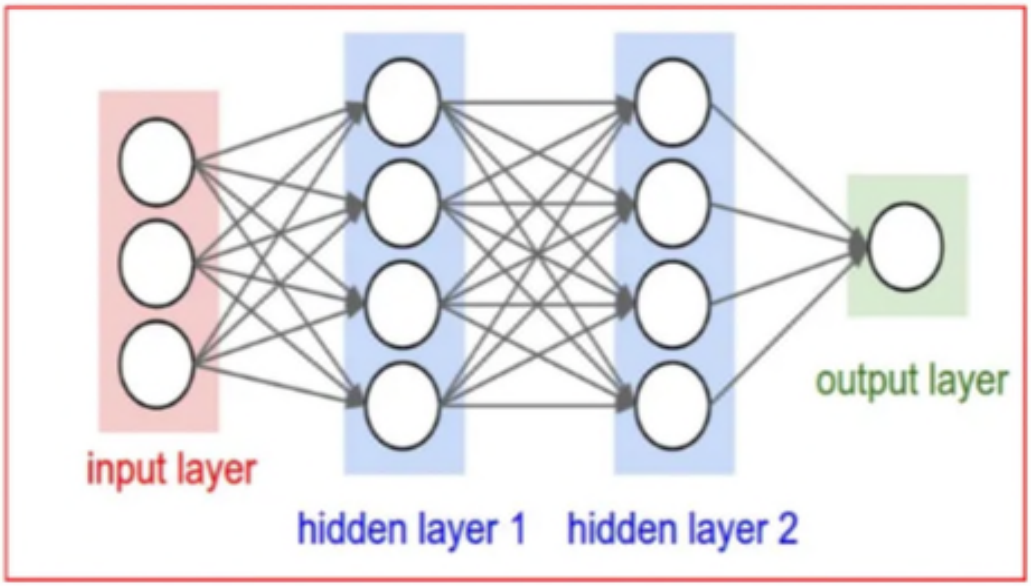


Рисунок 2.9 – Повнозв’язний рівень [23]

2.1.5 Метод навчання нейронної мережі

Найважливішою частиною у коректності роботи будь-якої нейронної мережі є її метод навчання. Для представленої архітектури нейронної мережі було обрано

найпопулярніший метод навчання зворотне поширення. Зворотне поширення представляє собою ітеративний метод навчання суть якого полягає у знаходженні градієнту кожного нейрона по відношенню до функції втрат, щоб визначити, наскільки вихідні дані вносять вклад в загальні втрати. В даному методі процес навчання відбувається з кінця, тобто поширення сигналів помилки йдуть від виходів мережі до її входів, з точки зору представленої архітектури від повністю зв’язаного шару до шару згортки [25].

Першим кроком в роботі даного методу необхідно визначити функцію втрат, яка використовується для оцінки виходу, отриманого з останнього рівня повнозв'язної мережі. Вибір функції втрат є важливим аспектом у побудові нейронної мережі, оскільки це великою мірою впливає на швидкість та якість навчання мережі, а також на її здатність оцінювати точність отриманих результатів. Значення функції втрат є основним критерієм для оцінки продуктивності мережі, оскільки вона дозволяє об'єднати всі переваги, недоліки та помилки мережі в одне числове значення, що характеризує мережу таким чином, що результати роботи для різних варіантів моделі можуть бути впорядковані за точністю. Вибір конкретної функції втрати повністю залежить від поставленої мети та завдання, яке планується виконати за допомогою мережі.

В залежності від типу задачі, для нейронних мереж використовуються два основних типи функцій втрат: функція перехресної ентропії та функція середньоквадратичної похибки. Поставлена задача представляє собою класифікаційний тип задач, так як потрібно визначити чи являється знайдена ознака у відеопотоці насильницькою. В задачах класифікації, використовуються ймовірності, що вхідні дані належать різним класам, для порівняння результатів мережі з визначеними класами. У таких задачах параметри мережі визначаються за принципом максимальної правдоподібності, щоб наблизити розподіл результатів до розподілу справжніх даних. Таким чином, критерієм помилки мережі є різниця між імовірностями належності вхідних даних до різних класів, отриманих мережею, та заданих у вибірці. Тому, в задачах класифікації з використанням згорткових мереж,

переважно використовується функція перехресної ентропії. Наведено формулу (7) для розрахунку втрат бінарної крос-ентропії, на рисунку 2.10 зображено виконання даної операції:

$$L(y, \hat{y}) = -\frac{1}{n}\sum_{i=1}^{n}(y_i)\log(\hat{y}_i) + (1 - y_i)\log(1 - \hat{y}_i) \quad (7)$$

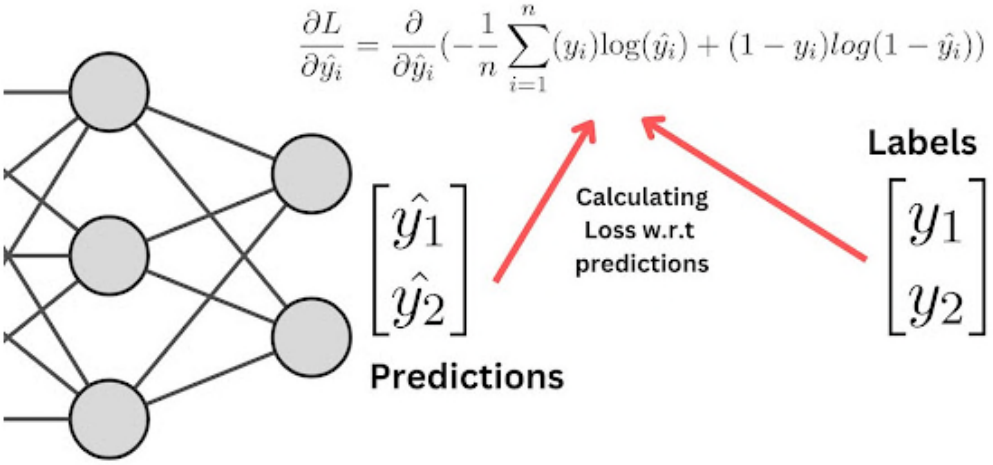


Рисунок 2.10 – Розрахунок втрат бінарної крос-ентропії [25]

У наступному кроці роботи методу зворотного поширення необхідно знайти градієнти втрат відносно ваг, вхідних даних та зміщення. Знаходження градієнтів є важливим аспектом у роботі нейронної мережі, так як вони забезпечують коректне оновлення вагових коефіцієнтів мережі під час процесу навчання. Перший градієнт який необхідно знайти це градієнт втрат відносно ваг. Даний градієнт вказує, наскільки зміни ваги впливають на функцію втрат, тим самим допомагає знайти градієнт втрат щодо ваг. Знайти градієнт втрат відносно ваг можна за допомогою формули (8), на рисунку 2.11 зображено виконання даної операції:

$$\frac{\partial L}{\partial \hat{y}_i} * x_j = -\frac{1}{n}\left(\frac{y_i}{\hat{y}_i} - \frac{1 - y_i}{1 - \hat{y}_i}\right) * x_j \quad (8)$$

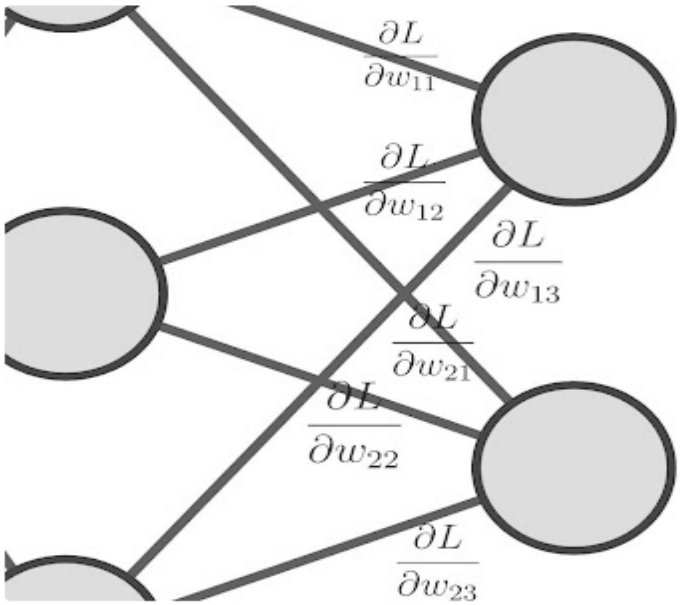


Рисунок 2.11 – Знаходження градієнту втрат відносно ваг [25]

Після знаходження градієнту втрат відносно ваг необхідно знайти градієнт втрат щодо вхідних даних, тобто, наскільки зміна вхідних даних впливає на втрати, нижче наведено формулу (9) для знаходження даного градієнту, на рисунку 2.12 зображено виконання даної операції:

$$\frac{\partial L}{\partial x_j} = -\frac{1}{n}\sum_{i=1}^{n}\left(\frac{y_i}{\hat{y}_i} - \frac{1-y_i}{1-\hat{y}_i}\right) * w_{ij} \tag{9}$$

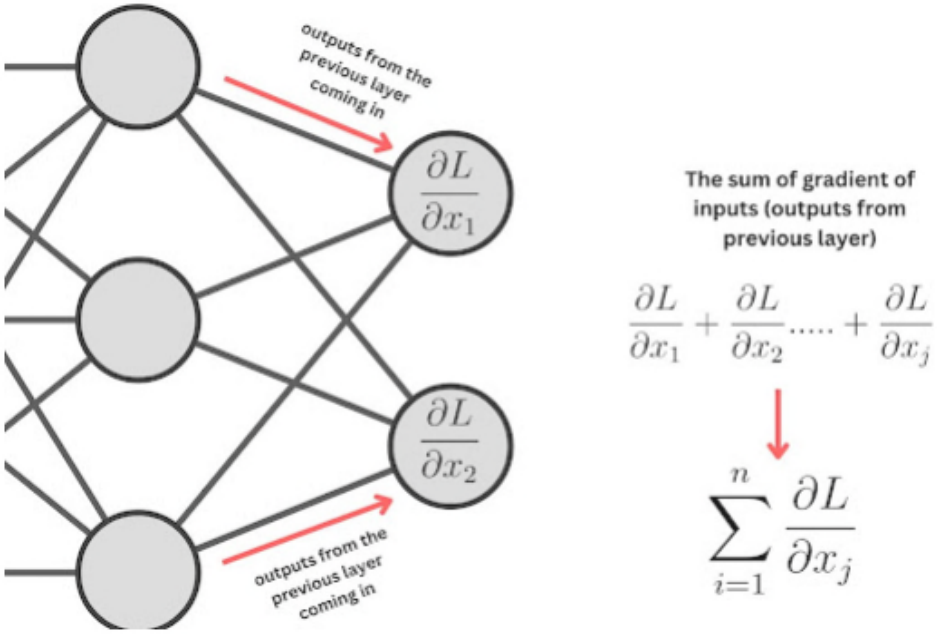


Рисунок 2.12 – Знаходження градієнту втрат відносно вхідних даних [25]

Останній градієнт, який потрібно знайти це градієнт втрат відносно зсуву. Даний градієнт можна знайти за наступною формулою (10):

$$\frac{\partial L}{\partial b_j} = \frac{\partial L}{\partial \hat{y}_i} * \frac{\partial \hat{y}_i}{\partial b_j} \tag{10}$$

Знаходження градієнту втрат відносно зсуву показує наскільки втрати змінюються по відношенню до виходів і наскільки змінюється вихід по відношенню до зсуву.

Після знаходження градієнтів наступає найважливіший крок у методі зворотного поширення – оновлення ваг і зміщення за допомогою градієнтів. Для оновлення ваг використовується формула (11):

$$w_{ij}^{new} = w_{ij}^{old} - \eta * \frac{\partial L}{\partial w_{ij}} = w_{ij}^{old} + \eta * \frac{1}{n}\left(\frac{y_i}{\hat{y}_i} - \frac{1 - y_i}{1 - \hat{y}_i}\right) * x_j, \tag{11}$$

де η – швидкість навчання,

w_{ij}^{new} – оновлені ваги,

w_{ij}^{old} – старі ваги,

$\frac{\partial L}{\partial w_{ij}}$ – градієнт втрат відносно ваг.

Формула (12) для оновлення зміщення:

$$b_j^{new} = b_j^{old} - \eta * \frac{\partial L}{\partial b_j} = b_j^{old} + \eta * \frac{1}{n}\left(\frac{y_i}{\hat{y}_i} - \frac{1 - y_i}{1 - \hat{y}_i}\right), \quad (12)$$

де η – швидкість навчання,

b_j^{new} – оновлене зміщення,

b_j^{old} – старе зміщення,

$\frac{\partial L}{\partial b_j}$ – градієнт втрат відносно зсуву.

На даному етапі завершено обчислення градієнтів та оновлення ваг повнозв'язного рівня.

Наступним кроком в роботі методу зворотного поширення є передача градієнтів вхідних даних, створених із повнозв'язного шару до шару максимального об'єднання. Але перед передачею даних з повнозв'язного шару до шару максимального об'єднання необхідно зробити інверсійний процес перетворення даних, тобто перетворити вектора повнозв'язного шару в матрицю відповідної розмірності шару максимального об'єднання. На рисунку 2.13 зображено виконання даної операції.

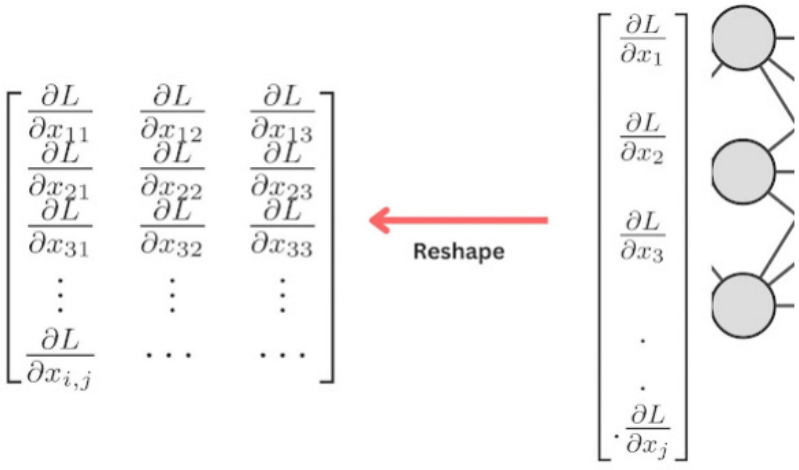


Рисунок 2.13 – Процес перетворення вектора повнозв'язного шару в матрицю відповідної розмірності шару максимального об'єднання [25]

При зворотному проході процес зміни форми можна визначити, спираючись на форму шарів об'єднання при прямому проході. Форма об'єднаних шарів визначатиме, яким чином одновимірний вектор градієнта з повнозв’язного шару повинен бути перетворений назад у градієнтну матрицю відповідної розмірності. При розгляді максимального об’єднання, важливо відзначити, що обчислення градієнта насправді не відбувається. Насправді, градієнт передається назад відповідно до функції Pooling, яка використовується. При максимальному об’єднанні вибирається максимальне значення з кожного вікна і передається у прямому проході. У зворотному поширенні градієнт, який відповідає цим максимальним значенням, передається назад, а решта значень вважається 0. Наведено формулу (13), яка відображає процес зворотного поширення на рівні максимального об’єднання:

$$\frac{\partial L}{\partial x_{ij}} = \frac{\partial L}{\partial \hat{y}_{ij}} * \delta_{ij,} \tag{13}$$

де $\delta_{ij} = 1$, якщо максимальне розташування пулу / 0, інше.

δ_{ij} є змінною перемикання або маскою, яка використовується для перемикання між значеннями 1 і 0 для різних позицій у матриці градієнта. Якщо значення дорівнює 1, то градієнт передається назад, а якщо значення дорівнює 0, то градієнт не передається. Значення 1 призначається на позиціях, де знайдено максимальні значення, а решта позицій мають значення 0.

Останній крок в роботі методу зворотного поширення є робота з згортковим шаром. Подібно до повнозв’язного шару необхідно знайти градієнти відносно ваг, вхідних даних і зміщення.

Першим етапом є знаходження градієнту втрат відносно ваг. Виконання даної операції наведено на рисунку 2.14, нижче наведено формулу (14) для виконання представленої операції:

$$\frac{\partial L}{\partial K_{ijk}} = \frac{\partial L}{\partial \hat{y}_{ij}} * \frac{\partial \hat{y}_{ij}}{\partial rot180\left(K_{ijk}\right)} = \frac{\partial L}{\partial \hat{y}_{ij}} * x_{ij} \tag{14}$$

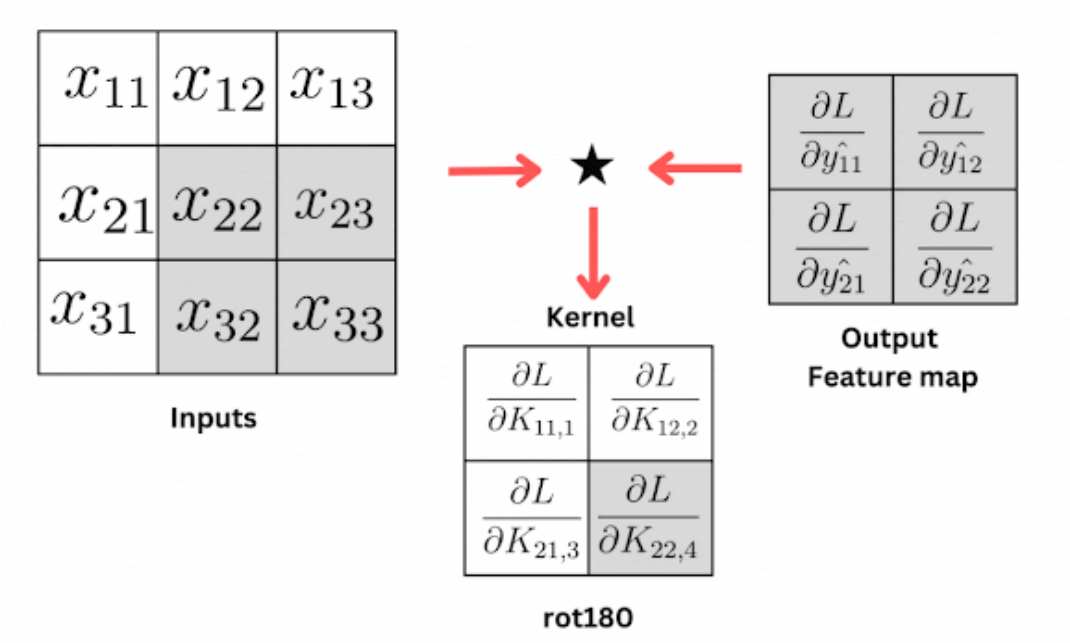


Рисунок 2.14 – Градієнт втрат відносно ваг [25]

Наступний етап є знаходження градієнту втрат відносно вхідних даних у шарі згортки. Виконання даної операції наведено на рисунку 2.15, наведено формулу (15) для виконання представленої операції:

$$\frac{\partial L}{\partial x_{ij}} = \sum_{i=1}^{m}\sum_{j=1}^{n} \frac{\partial L}{\partial \hat{y}_{ij}} * rot180\left(K_{ijk}\right) \tag{15}$$

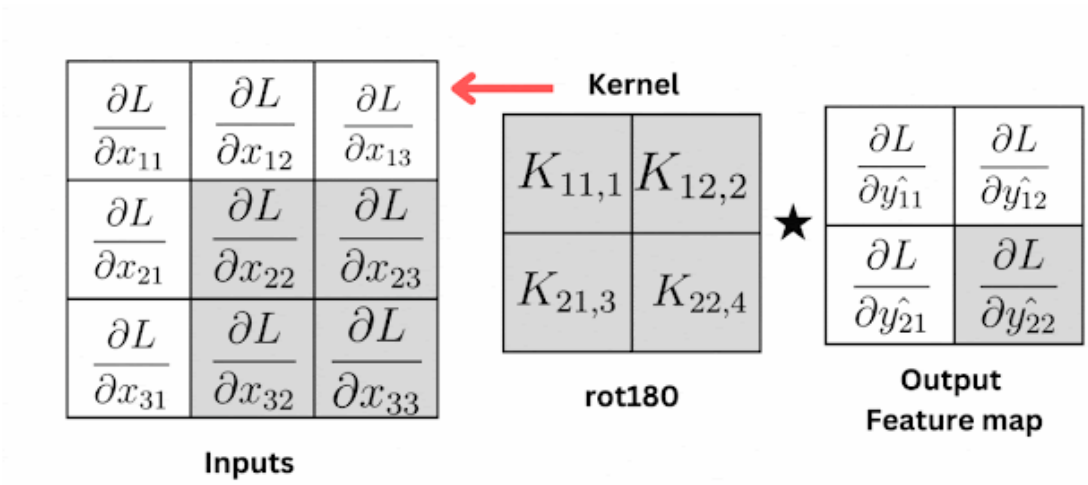


Рисунок 2.15 – Градієнт втрат відносно вхідних даних у шарі згортки [25]

Градієнт втрат відносно вхідних даних визначається подвійним підсумовуванням змін у втратах відносно вихідних даних, отриманих після крос-кореляції та повороту ядра на 180 градусів.

Останній градієнт який потрібно знайти це градієнт втрат відносно зміщення. Нижче представлено формулу (16) для знаходження даного градієнту втрат:

$$\frac{\partial L}{\partial b_j} = \sum_{j=1}^{m} \frac{\partial L}{\partial \hat{y}_{ij}} \tag{16}$$

Завершальним кроком в роботі методу зворотного поширення є оновлення ядра та зміщення за допомогою методу градієнтного спуску. Нижче представлено формулу (17) для оновлення ядра:

$$K_{ijk}^{new} = K_{ijk}^{old} - \eta * \frac{\partial L}{\partial K_{ijk}} \tag{17}$$

Формула (18) для оновлення зміщення:

$$b_j^{new} = b_j^{old} - \eta * \frac{\partial L}{\partial b_j} \tag{18}$$

2.1.6 Метод тонкого налаштування нейронної мережі

Особливістю запропонованого методу є можливість працювати з відеопотоком у реальному часі (пряма відеотрансляція). Дана можливість досягається за рахунок використання методу тонкого налаштування (fine-tuning) згортковою нейронною мережею. Fine-tuning є процесом використання попередньо навченої моделі машинного навчання і подальшого навчання цієї моделі на нових даних для покращення її точності на конкретній задачі. Основна ідея застосування fine-tuning полягає в тому, що попередньо навчена модель вже має досвід вирішення загальних завдань, у випадку поставленої задачі це виявлення ознак насильства на статичних відеоматеріалах. Цей досвід може бути використаний, щоб покращити точність

моделі на задачі виявлення зовнішніх проявів насильства у прямій відеотрансляції (динамічний відеоматеріал).

У контексті запропонованого методу тонке налаштування згорткової нейронної мережі можна поділити на декілька етапів:

Етап 1 – Замороження ваг та параметрів

Першим кроком необхідно заморозити всі ваги та параметри попередньо навченої моделі, крім останнього шару. Це дозволить зберегти вивчені ознаки, які були навчені на задачі виявлення ознак насильства у статичних відеоматеріалах, але дозволить моделі навчати нові шари для розпізнавання насильства у прямій відеотрансляції.

Етап 2 – Додати новий шар

Другим кроком потрібно додати новий повнозв’язний (full-connected) шар до попередньо навченої моделі для адаптації до нової задачі. Додавання нового повнозв'язного шару дозволяє провести остаточну обробку признакового вектора, тим самим допомогти моделі вчити більш складні та абстрактні залежності в даних.

Етап 3 – Навчання моделі

Третім кроком є навчання моделі на неперервному потоку даних з мультимедійних платформ для онлайн трансляцій, використовуючи метод зворотного поширення помилки. Під час навчання модель буде налаштовувати ваги та параметри нового шару, щоб вона була краще пристосована до конкретної задачі, а саме виявлення зовнішніх проявів насильства у прямій відеотрансляції.

2.2 Визначення характеру відео з використанням класифікатора SVM

SVM (метод опорних векторів) є алгоритмом машинного навчання, який використовується для класифікації та регресійного аналізу даних. Він базується на моделях з керованим навчанням та використовує пов'язані алгоритми для ефективного навчання та прогнозування [24]. Основна ціль SVM як класифікатора полягає в пошуку роздільної гіперплощини $w_1x_1 + w_2x_2 + \ldots + w_nx_n + w_0 = 0$ в

просторі R^n, яка оптимальним чином розділяє два класи, у випадку поставленої задачі це насильницький або не насильницький відеоматеріал.

Робота алгоритму SVM полягає в наступному: першим кроком SVM тренується на об'єктах з навчальної вибірки, котрим заздалегідь відомі мітки класів. Далі навчений алгоритм передбачає мітку класу для кожного об'єкта з тестової вибірки. Мітки класів як правило в даному алгоритмі приймають значення $Y = \{-1,1\}$. Об’єкт є вектором з N ознаками $x = (x_1, x_2, \dots, x_n)$ в просторі R^n. При навчанні алгоритм повинен побудувати функцію $F(x) = y$, яка приймає аргумент x – об’єкт із простору R^n і видає мітку класу y.

Перетворення F об'єкта x у мітку класу Y може бути виражено як $F(x) = \sinh(w^T x - b)$. Функція sinh використовується для визначення мітки на основі знаку виразу $w^T x - b$, $w = (w_1, w_2, \dots, w_n)$, $b = -w_0$. Після налаштування ваг алгоритму w і b (навчання), всі об'єкти, які потрапляють по одну сторону від побудованої гіперплощини, будуть передбачатися як насильницький клас, тоді як об'єкти, що потрапляють по іншу сторону, будуть передбачатися як не насильницький.

У методі опорних векторів (SVM) ваги w і b налаштовуються таким чином, щоб забезпечити якнайбільший зазор між роздільною гіперплощиною і об'єктами класів, які знаходяться найближче до неї. Це означає, що алгоритм намагається максимізувати відстань між гіперплощиною і найближчими об'єктами класів. Об'єкти, які розташовані найближче до роздільної гіперплощини і впливають на налаштування ваг, називаються опорними векторами. Це означає, що ці об'єкти є критичними для визначення роздільної гіперплощини і впливають на класифікацію нових об'єктів. На рисунку 2.16 зображено алгоритм SVM.

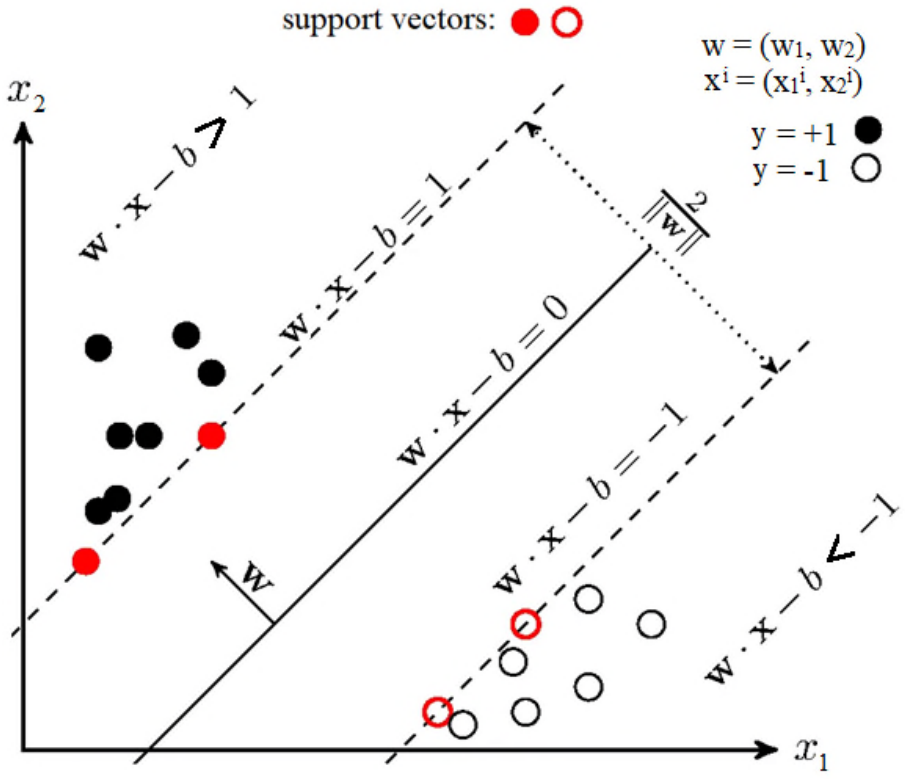


Рисунок 2.16 – Алгоритм роботи SVM [24]

Найважливішим елементом в роботі даного алгоритму є коректне налаштування ваг. Для налаштування ваг необхідно забезпечити максимальну віддаленість роздільної гіперплощини від точок вибірки, тобто потрібно максимізувати ширину смуги. Вектор w є вектором нормалі до роздільної гіперплощини. Для позначення скалярного добутку двох векторів використовується вираз (a, b) або $a^T b$. Необхідно знайти проекцію вектора, який має опорні вектори різних класів як свої кінці на вектор w, ця проекція буде відображати ширину розділяючої смуги.

Відступ об'єкта x від межі класів визначається як величина $M = y(w^T x - b)$. Алгоритм вважається помилковим на об'єкті тільки в тому випадку, коли відступ M є від'ємним (коли y та $w^T x - b$ мають протилежні знаки). Якщо $M \in (0,1)$, то об'єкт потрапляє всередину роздільної смуги. Якщо $M > 1$, то об'єкт x класифікується правильно і знаходиться на певній відстані від роздільної смуги. Таким чином, алгоритм правильно класифікуватиме об'єкти, якщо виконується умова (19):

$$y(w^T x - b) \geq 1 \quad (19)$$

Опираючись на наведене рівняння (19) можна знайти функцію втрат SVM, рівняння (20):

$$Q = \max(0,1 - yw^T x) + \frac{\alpha(w^T w)}{2}. \quad (20)$$

Отримавши функцію втрат можна вивести правила оновлення ваг використовуючи градієнтний спуск, рівняння (21):

$$w = w - \eta \Delta Q,$$

де $\eta -$ крок спуску.

Звідси,

$$\Delta Q = \begin{cases} \alpha w - yx, \text{якщо } yw^T x < 1 \\ \alpha w, \text{якщо } yw^T x \geq 1 \end{cases} \quad (21)$$

На рисунку 2.17 зображено правила налаштування ваг в SVM.

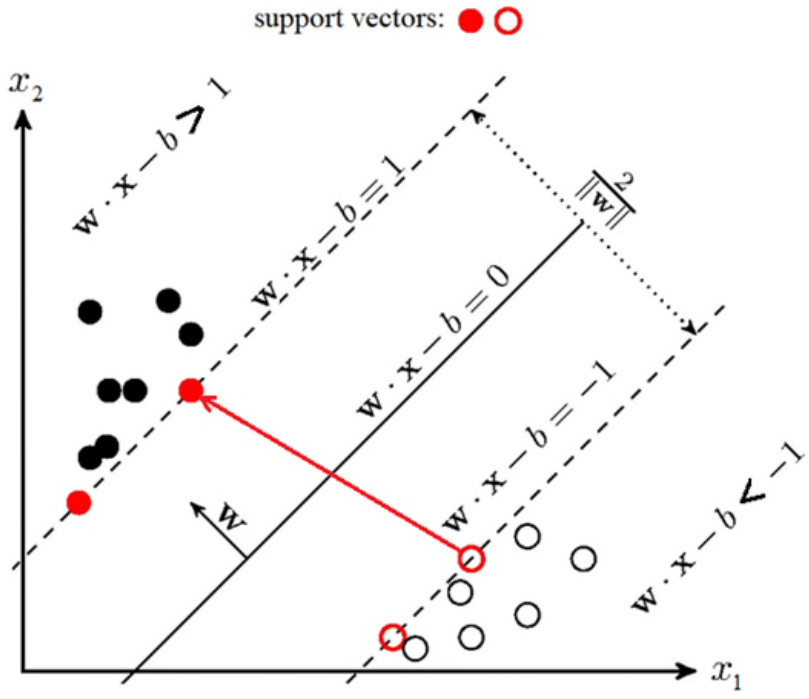


Рисунок 2.17 – Правила налаштування ваг в SVM [24]

Важливим аспектом у роботі класифікатора SVM є визначення порогу (значення від якого залежить визначення належності вхідних даних до відповідного класу) для задачі виявлення зовнішніх проявів насильства у відеопотоці, оскільки він впливає на баланс між чутливістю (здатність виявляти всі випадки насильства) та точністю (правильність визначення насильства). Тобто, чим більше значення порогу тим більша буде точність, але буде менша чутливість (є ймовірність, що не врахує грубі штовханини), і навпаки менше значення порогу супроводжує більшу чутливість, але меншу точність (є ймовірність, що звичайні обійми можуть враховуватися як насильницькі дії).

Вибір порогу для класифікатора напряму залежить від набору даних з яким він працюватиме, очікуваних результатів (тобто визначення пріоритетності щодо точності та чутливості) та експериментів зі самим значенням з метою знаходження оптимального значення.

В контексті запропонованого методу для класифікатора SVM обрано діапазон визначення оцінки від 0 до 1 (тобто від 0% до 100%) та поріг 0.5 (оцінка менше рівне 0.5 свідчить про відсутність насильства, значення більше 0.5 – насильство (чим ближче значення до 1 тим очевидніший акт насильства)). Значення 0.5 в діапазоні визначення оцінки від 0 до 1 свідчить про баланс між чутливістю та точністю. Дане значення обрано з врахуванням набору даних на якому навчалася згорткова нейронна мережа. Набори даних включали в себе не тільки прямі акти насильства (типу прямої боротьби), а й дії типу штовханин, обнімань тощо. Також враховано пріоритетність між чутливістю та точністю, значення 0.5 забезпечить коректне визначення насильства не лише в контексті прямої бійки, але і грубих штовханин і при цьому зможе коректно врахувати такий контакт (ненасильницький) між людьми як звичайні обійми.

Висновки до розділу 2

В другому розділі розроблено метод для виявлення зовнішніх проявів насильства у відеопотоці нейромережевими засобами. В якості нейромережевого засобу було розглянуто згорткову нейрону мережу (CNN). Наведено архітектуру даної мережі з описами кожного шару та рівнів, а також представлено математичні формули для відповідних операцій.

Розглянуто метод навчання для представленої нейронної мережі. В якості методу навчання було обрано зворотне поширення помилки. Описано потік роботи даного методу, а також математичні формули для операцій в даному методі.

Розглянуто алгоритм SVM, який виступає в якості класифікатора з цілю визначення характеру відео, тобто чи представлений відеоматеріал несе насильницький характер чи ні у відсотковому відношенні.

Розглянуто метод тонкого налаштування (fine-tuning) для згорткової нейронної мережі. Ціль використання даного методу полягає у адаптації запропонованого методу виявляти ознаки насильства у прямій відеотрансляції.

РОЗДІЛ 3

Програмна реалізація методу виявлення зовнішніх проявів насильства у відеопотоці нейромережевими засобами

3.1 Вибір платформи, технологій та бібліотек

Для реалізації методу виявлення зовнішніх проявів насильства у відеопотоці нейромережевими засобами необхідно розробити додаток, який дозволить маніпулювати запропонованим методом і відповідно тестувати його. В якості середовища, де буде функціонувати додаток обрано веб-середовище, так як воно дозволяє легко обробляти великі обсяги відеоданих через можливість працювати з розподіленими системами та використовувати хмарні рішення для обробки даних, що забезпечує високу продуктивність та швидкість обробки. Додатки, побудовані у веб-середовищі, можуть легко розповсюджуватися та оновлюватися. Це дозволяє швидко впроваджувати оновлені моделі виявлення насильства та забезпечувати користувачам оновлення без необхідності переустановлення додатку [26].

Основні характеристики веб-середовища:

– *Веб-сервер.* Веб-середовище базується на використанні веб-сервера, який є основою веб-архітектури. Веб-сервер приймає HTTP-запити від клієнтів і передає їх до веб-додатку для обробки. Він також надсилає HTTP-відповіді від веб-додатку до клієнта.

– *Мови програмування.* Дане середовище підтримує різні мови програмування, такі як Javascript, PHP, Python, Ruby, Java тощо. Це дозволяє розробникам використовувати мову, з якою вони найбільш знайомі.

– *Розширення та бібліотеки.* Веб-середовище надає розширення та бібліотеки, що допомагають розробникам виконувати різні завдання. Наприклад, бібліотеки можуть бути використані для роботи з базами даних, криптографією, роботою з файлами тощо.

– *Безпека.* Середовище забезпечує безпеку веб-додатків шляхом застосування різних методів, таких як автентифікація, авторизація, шифрування тощо. Воно також може забезпечувати захист від атак, таких як внесення виразів, перетин сайту, SQL-ін'єкція тощо.

– *Керування сесіями.* Веб-середовище забезпечує механізми для керування сесіями. Це дозволяє зберігати стан між запитами клієнта і забезпечувати інтерактивність веб-додатків, таку як увійти в систему, корзина покупок тощо.

– *Логування та моніторинг.* Дане середовище надає можливості логування та моніторингу, які допомагають розробникам зрозуміти, як працює їх веб-додаток, виявляти помилки та проблеми продуктивності.

– *Розширюваність.* Веб-середовище часто має можливості для розширення, що дозволяє додавати функціональність через модулі, плагіни або додатки. Це дозволяє розробникам налаштовувати середовище під свої потреби [26].

Програмну реалізацію запропонованого методу можна представити у вигляді веб-додатку, де ядром та ключовим елементом є метод виявлення зовнішніх проявів насильства у відеопотоці нейромережевими засобами.

Веб-додатки є програмами, які працюють у веб-браузері і доступні користувачам через Інтернет. Вони використовуються для виконання різних завдань, від роботи з електронною поштою до онлайн-покупок і соціальних мереж. Особливості веб-додатків включають наступне:

– *Доступність.* Оскільки веб-додатки працюють у веб-браузері, їх можна використовувати на будь-якому пристрої з доступом до Інтернету, такому як комп'ютери, смартфони або планшети. Користувачі можуть легко отримати доступ до веб-додатків без необхідності встановлювати додаткове програмне забезпечення.

– *Кроссплатформеність.* Веб-додатки розробляються з використанням веб-технологій, таких як HTML, CSS і JavaScript, що дозволяє їм працювати на різних операційних системах, таких як Windows, macOS або Linux. Це полегшує розповсюдження та використання додатків на різноманітних платформах.

– *Оновлення.* Одна з переваг веб-додатків полягає в тому, що розробники можуть оновлювати їх безпосередньо на сервері. Користувачі отримують оновлення автоматично при наступному відвідуванні додатку, що забезпечує постійне покращення та виправлення помилок.

– *Масштабованість.* Веб-додатки можуть працювати з низьким або високим обсягом трафіку, залежно від потреб користувачів. Вони можуть бути легко масштабовані для обробки великого числа користувачів або підвищення продуктивності.

– *Зручний доступ до даних.* Веб-додатки можуть зберігати та обробляти дані на сервері, що дозволяє користувачам отримувати доступ до своїх даних з будь-якого пристрою з Інтернетом. Крім того, веб-додатки можуть інтегруватися з іншими системами та послугами, що робить їх більш універсальними.

– *Безпека.* Розробники веб-додатків забезпечують безпеку своїх додатків шляхом застосування різних заходів безпеки, таких як шифрування даних, автентифікація користувачів та захист від хакерських атак. Використання веб-додатків також дозволяє забезпечити оновлення безпеки безпосередньо на сервері [26].

Сьогодні сучасні веб-додатки розробляються у вигляді 2 складових: frontend і backend.

Для реалізації frontend частин обрано фреймворк Electron.js. Даний фреймворк є потужним інструментом для розробки десктопних та веб-додатків з використанням веб-технологій, таких як HTML, CSS і JavaScript. Цей фреймворк надає можливість розширити функціональність застосунків за допомогою системи доповнень. Присутня можливість використовувати модулі Node.js та розширений API для створення нативних діалогів, інтеграції з іншими застосунками, створення контекстних меню, виведення повідомлень, маніпуляції вікнами та взаємодії з підсистемами Chromium [27]. Основні особливості Electron.js:

– *Кроссплатформеність*. Electron.js дозволяє розробникам створювати додатки, які працюють на різних операційних системах без необхідності писати

окремий код для кожної платформи. Це дозволяє ефективно використовувати ресурси та забезпечує єдиний досвід користувача на різних платформах.

– *Використання веб-технологій.* Фреймворк використовує веб-технології, такі як HTML, CSS і JavaScript, для створення інтерфейсу користувача та логіки додатків. Це означає, що веб-розробники можуть використовувати свої вміння та знання для створення настільних додатків.

– *Можливості Node.js.* Electron.js базується на Node.js, що дозволяє розробникам використовувати Node.js-модулі та бібліотеки в своїх додатках. Це розширює можливості розробки та дозволяє використовувати потужні інструменти та функціональність Node.js.

– *Можливості доступу до системи.* Технологія надає розробникам можливість отримувати доступ до системних ресурсів та функцій, таких як файлова система, мережа, операційна система та багато іншого. Це дозволяє створювати потужні додатки з розширеними можливостями [27].

В контексті запропонованого методу даний фреймворк забезпечує можливість методу працювати повноцінно у веб-середовищі, а також при необхідності є можливість легко розгорнути додаток на локальній машині.

В якості інструменту для реалізації компонентної розмітки додатку та його стилізації обрано фреймворк Bootstrap 4. Bootstrap 4 є безкоштовним фреймворк для розробки веб-сайтів, який допомагає створювати респонсивні та мобільно-дружні сайти. Основна мета даного фреймворку – спростити процес розробки, надаючи готові HTML та CSS шаблони, які можна використовувати для створення різних компонентів і макетів. Основні особливості Bootstrap 4:

– *Респонсивний дизайн.* Bootstrap 4 надає гнучкі та адаптивні класи, які дозволяють легко налаштовувати вигляд сайту для різних розмірів екранів.

– *Готові компоненти.* Фреймворк містить велику кількість готових компонентів, таких як кнопки, форми, навігація, каруселі, модальні вікна та багато інших. Це дозволяє швидко створювати функціональні елементи на сайті.

– *Сітка.* Bootstrap 4 має потужну систему сітки, яка допомагає організувати контент на сторінці. Сітка дозволяє легко розміщувати елементи на різних розмірах екранів і забезпечує їх адаптивність.

– *Теми та налаштування.* Bootstrap 4 надає можливість налаштувати вигляд сайту, використовуючи різні теми та налаштування. Це дозволяє створювати унікальний дизайн, відповідний вимогам проекту.

– *Підтримка браузерів.* Bootstrap 4 підтримує останні версії всіх популярних браузерів, забезпечуючи сумісність із широким колом пристроїв та платформ [28].

Для реалізації backend частини обрано фреймворк Django. Django є високорівневим фреймворк для розробки веб-додатків на мові програмування Python. Він надає потужні інструменти для швидкої і ефективної розробки веб-додатків, забезпечуючи готову основу для роботи з базами даних, обробки URL-адрес, шаблонів, безпеки та багато іншого [29]. В контексті запропонованого методу даний фреймворк використовується для розробки веб-додатку, який надає можливість завантажувати відеофайли та обробляти їх за допомогою навчених моделей. Основні характеристики фреймворку Django:

– *Модульність.* Django має велику кількість готових модулів, які допомагають в розробці різних функціональностей, таких як автентифікація користувачів, робота з базами даних, кешування, обробка форм і багато іншого.

– *Адміністративний інтерфейс.* Django надає готовий адміністративний інтерфейс, який дозволяє легко керувати даними в базі даних без необхідності написання власного коду.

– *MVC архітектура.* Django використовує MVC-архітектуру для розділення логіки додатку на моделі (дані), види (представлення) та контролери (логіка). Це дозволяє розробникам легко організовувати код та підтримувати його.

– *ORM.* Django надає ORM, що дозволяє розробникам працювати з базою даних, використовуючи об'єктно-орієнтований підхід. Це означає, що вони можуть

взаємодіяти з базою даних, використовуючи об'єкти та методи, що забезпечує більш зрозумілий та зручний спосіб роботи з даними.

– *Швидкодія.* Django працює швидко завдяки своїм оптимізованим алгоритмам та кешуванню, що дозволяє збільшити продуктивність веб-додатків. Він також надає можливість масштабування та оптимізації для обробки великого обсягу трафіку.

– *Безпека.* Django має вбудовані заходи безпеки, такі як захист від хакерських атак, захист від перехоплення сесій, захист від введення шкідливого коду і багато іншого.

– *Спільнота та документація.* Django має активну спільноту розробників, яка надає підтримку та відповідає на питання. Також існує велика кількість документації, ресурсів та підручників, які допомагають розробникам вивчити та використовувати Django ефективно.

– *Розширюваність.* Django дозволяє розширювати функціональність за допомогою власних модулів і плагінів, що дозволяє створювати унікальні веб-додатки [30].

Для реалізації запропонованого методу обрано мову програмування Python. Python є мовою програмування, яка широко використовується в області розробки нейромереж. Python надає потужні бібліотеки та фреймворки для розробки нейромережевих моделей, що дозволяють впроваджувати метод виявлення зовнішніх проявів насильства у відеопотоці [31].

Основні характеристики Python:

– *Заснований на об'єктах.* Python підтримує об'єктно-орієнтовану парадигму програмування, що дозволяє створювати класи, об'єкти та використовувати спадкування та поліморфізм.

– *Зрозумілий синтаксис.* Синтаксис Python дуже читабельний та зрозумілий, що полегшує розробку та зрозуміння коду.

– *Багатий набір бібліотек.* Python має велику кількість стандартних бібліотек, які допомагають в різних аспектах програмування, таких як робота з файлами, мережеве програмування, обробка рядків, математичні обчислення та багато іншого.

– *Переносимість.* Python підтримує переносимість між різними операційними системами, такими як Windows, macOS та Linux. Розроблені на Python програми можуть працювати на будь-якій платформі, що робить мову дуже гнучкою та універсальною.

– *Широке застосування.* Python використовується в багатьох галузях, таких як веб-розробка, наукові дослідження, штучний інтелект, аналітика даних, автоматизація та багато іншого. Його використання поширене у великих компаніях та стартапах через його простоту, продуктивність та екосистему.

– *Велика спільнота розробників.* Python має велику та активну спільноту розробників, яка надає підтримку, документацію та різні ресурси для вивчення та використання мови. Це дозволяє розробникам знайти відповіді на свої питання, отримати поради та підтримку в процесі роботи.

– *Підтримка різних платформ.* Python може працювати на різних операційних системах, включаючи Windows, macOS, Linux/UNIX та Android [32].

Основні бібліотеки Python, які використовуються для реалізації запропонованого методу:

– *TensorFlow.* TensorFlow є високопродуктивною бібліотекою для чисельних обчислень та машинного навчання. Вона розроблена для ефективної реалізації різних типів нейронних мереж та моделей глибокого навчання. Дана бібліотека надає набір готових алгоритмів чисельних обчислень, які реалізовані через графи потоків даних, що є іншими словами, структурою, де вузли представляють математичні операції або точки вводу/виводу, а ребра графу відображають багатовимірні масиви даних, відомі як тензори, які переміщуються між цими вузлами. Вузли можуть бути призначені для обчислювальних пристроїв і працювати асинхронно, виконуючи операції паралельно

над відповідними тензорами. Цей підхід дозволяє організувати одночасну роботу вузлів у нейронних мережах, подібно до одночасної активації нейронів у мозку [33].

– *Keras.* Keras є високорівневим інтерфейсом для розробки і навчання глибоких нейронних мереж. Він дозволяє легко створювати, навчати та експериментувати з нейронними мережами, приховуючи деталі нижчого рівня, такі як операції над тензорами і графи обчислень. Keras спрощує процес розробки нейронних мереж, роблячи його більш доступним для дослідників та розробників. Keras дозволяє швидко прототипувати моделі, використовуючи простий та інтуїтивний API, і він підтримує різні типи нейронних мереж, включаючи звичайні нейронні мережі, згорткові нейронні мережі (CNN), рекурентні нейронні мережі (RNN) та багатошарові перцептрони (MLP) [34].

– *OpenCV.* OpenCV є відкритою бібліотекою, яка надає набір інструментів та бібліотек для обробки та аналізу зображень і відео. OpenCV розробляється активною спільнотою та використовується в різних галузях, включаючи комп'ютерне бачення, розпізнавання образів, робототехніку, андроїд розробку, і багато інших. Вона має вбудовані функції для виявлення об'єктів, розпізнавання руху та аналізу відеопотоку. За допомогою OpenCV можна виконувати базовий аналіз відео, включаючи виявлення насильства на основі певних характеристик [35].

– *NumPy.* NumPy є бібліотекою для мови програмування Python, яка надає підтримку для великих масивів та матриць числових даних, разом із високорівневими математичними функціями для роботи з цими даними. NumPy є важливою складовою в екосистемі наукових обчислень в Python і використовується для розв'язання різноманітних завдань, таких як обробка даних, аналіз, статистика, машинне навчання та багато інших обчислювальних завдань [36].

– *Scikit-learn.* Scikit-learn є бібліотека для мови програмування Python, яка спеціалізується на машинному навчанні та аналізі даних. Вона надає широкий набір інструментів для навчання моделей машинного навчання, включаючи класифікацію, регресію, кластеризацію, метод опорних векторів (SVM), випадковий ліс, наївний

Баєсівський класифікатор, метод k-середніх, вимірювання якості моделей, а також підготовку та обробку даних [37].

– *Matplotlib.* Matplotlib є бібліотекою для мови програмування Python, яка використовується для створення візуалізацій та графіків. Вона надає інструменти для побудови різноманітних типів графіків, включаючи лінійні графіки, діаграми розподілу, стовпчикові діаграми, кругові діаграми, контурні графіки, графіки розсіювання та багато інших. Matplotlib дозволяє створювати візуальні представлення даних для аналізу та відображення результатів досліджень [38].

– *OpenVINO.* OpenVINO є набором інструментів, який розробники та аналітики даних можуть використовувати для розробки високопродуктивних рішень у відеосистемах. Цей набір інструментів є відкритим і безкоштовним і допомагає прискорити процес розробки шляхом надання підтримки для комп'ютерного зору. OpenVINO підтримує широкий спектр рішень для комп'ютерного зору і дозволяє оптимізувати розгортання моделей глибокого навчання. Крім того, він забезпечує просте виконання розроблених рішень на різних платформах, що базуються на технологіях Intel. Це дозволяє розробникам ефективно використовувати можливості глибокого навчання для візуального аналізу в різних додатках та системах [39].

3.2 Розробка прикладних компонентів додатку запропонованого методу

Для запропонованого методу реалізовано веб-додаток. Точкою входу будь-якого веб-додатку є головна сторінка, яка відображає основну ідею програмного продукту для користувачів. На рисунку 3.1 зображено головну сторінку веб-додатку методу виявлення зовнішніх проявів насильства у відеопотоці.

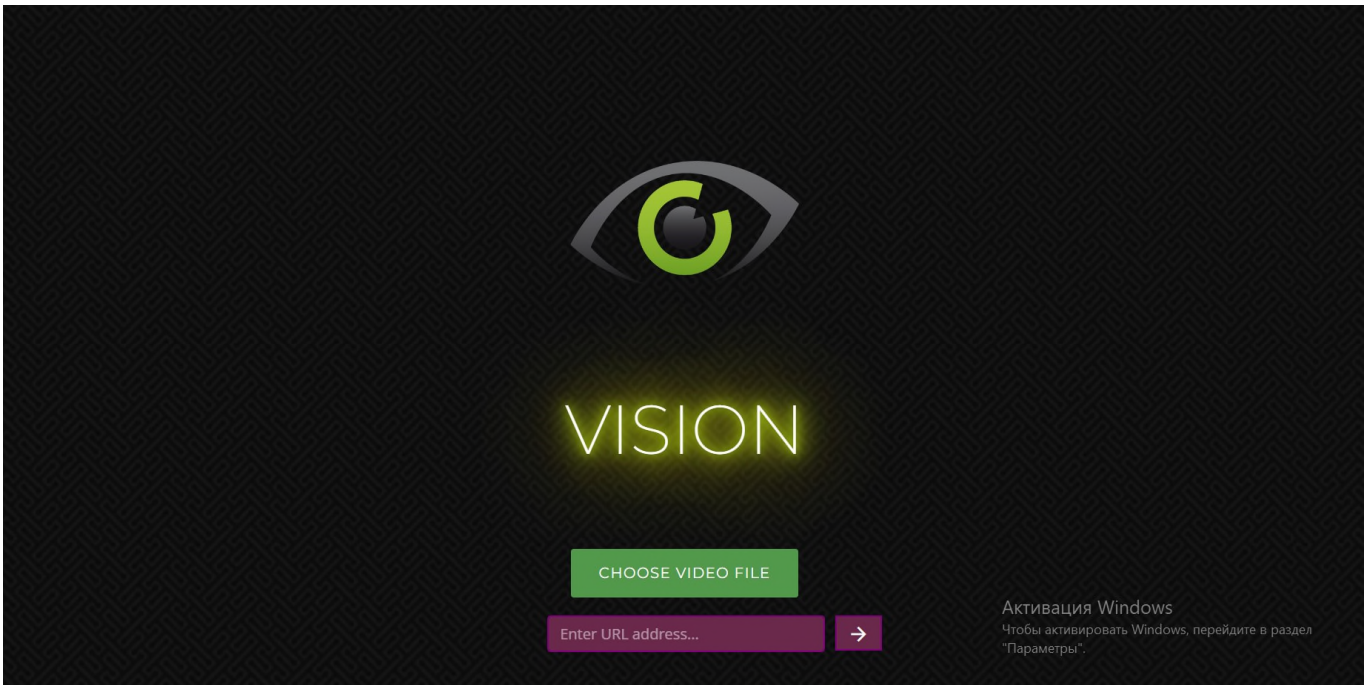


Рисунок 3.1 – Головна сторінка веб-додатку методу виявлення зовнішніх проявів насильства у відеопотоці

Для завантаження відеоматеріалу по якому здійснюватиметься процес виявлення зовнішніх проявів насильства використовується кнопка «CHOOSE VIDEO FILE». При натисканні даної кнопки користувачу необхідно вибрати відеоматеріал у форматі «.mp4» після чого додаток автоматично відправить дане відео на сервер для обробки.

На сервері відповідному методу подається запит, де він перевіряє тип вхідного запиту, токен валідації, а також наявність відеоматеріалу. Після проходження перевірок метод направляє вхідний відеоматеріал до нейронної мережі. На рисунку 3.2 зображено кнопку «CHOOSE VIDEO FILE» для завантаження відео.

Рисунок 3.2 – Кнопка CHOOSE VIDEO FILE» для завантаження відеоматеріалу

Для завантаження прямої відеотрансляції з метою виявлення зовнішніх проявів насильства використовується поле «Enter URL address». У дане поле вводиться url адреса прямої трансляції або онлайн відеокамери. Після вводу адреси користувачу потрібно натиснути на кнопку відправки після чого інформаційна система автоматично відправить інформацію про пряму трансляцію на сервер для обробки.

На сервері відповідному методу подається запит де він перевіряє тип вхідного запиту, токен валідації, а також наявність url адреси прямої трансляції. Після проходження перевірок метод використовуючи бібліотеку «OpenCV» встановлює з'єднання з прямою трансляцією по вхідній адресі і «захоплює» відеокадри трансляції, після чого направляє отриманий відеоматеріал до нейронної мережі. На рисунку 3.3 зображено кнопку поле «Enter URL address» для завантаження прямої трансляції.

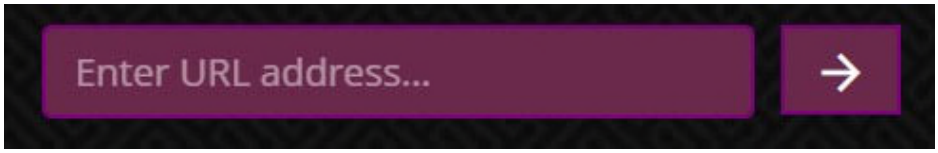


Рисунок 3.3 – Поле для введення URL адреси прямої трансляції

Перед тим як згорткова нейронна мережа розпочне процес вилучення ознак із вхідного відеоматеріалу запропонований метод вилучає із відео кадри та підготовлює їх до обробки нейронною мережею. Процес вилучення кадрів із вхідного відеоматеріалу здійснюється за допомогою методу «Extractor» та бібліотеки «OpenCV», після чого відбувається процес підготовки вилучених кадрів до обробки нейронною мережею.

Після підготовки вилучених кадрів із вхідного відеоматеріалу в роботу вступає згорткова нейронна мережа з метою вилучення ознак. Для вилучення ознак згорткова нейронна мережа здійснює наступні операції у декілька етапів: згортки, максимальне об'єднання, а також робота повнозв'язного рівня.

Завершальним етапом у визначенні насильства у відеопотоці є робота методу опорних векторів (SVM). Даний метод на основі вилучених ознак нейронною

мережею видає оцінку, яка є ймовірністю того, що вхідні дані належать до певного класу (насильницького або не насильницького).

Визначивши оцінку для вхідного відеоматеріалу, запропонований метод відображає дану оцінку у відеопотоці. Якщо оцінка має значення більше 0.5 то це означає, що у відеоматеріалі відбуваються дії які несуть у собі певний насильницький характер. В цьому випадку у відеопотоці відображається червоний прямокутник у якому вказано текст «Suspicious» та оцінка. В іншому випадку, якщо значення оцінки становить менше рівне 0.5, то у відеопотоці буде відображатися зелений прямокутник із текстом «Peaceful» та відповідна оцінка.

На рисунку 3.4 зображено сценарій у якому відеоматеріал несе насильницький характер, тобто оцінка більше 0.5, на рисунку 3.5 зображено сценарій у якому відсутнє насильство.

Рисунок 3.4 – Сценарій у якому відеоматеріал несе насильницький характер

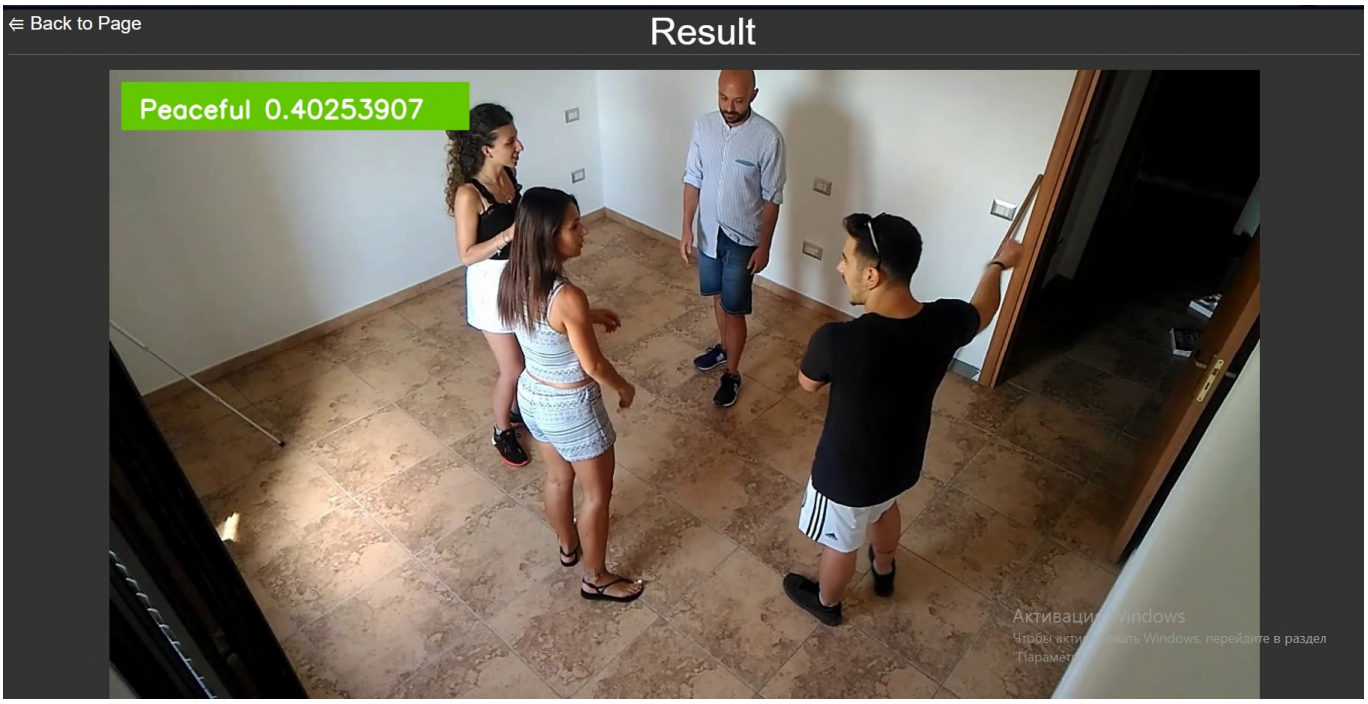


Рисунок 3.5 – Сценарій при якому у відеоматеріалі відсутнє насильство

3.3 Оптимізація методу виявлення зовнішні прояві насильства у відеопотоці нейромережевими засобами

У більшості випадків застосуванням глибокого навчання моделей штучного інтелекту потрібна оптимізація для ефективного використання обчислювальних ресурсів, щоб досягти високої продуктивності.

Оптимізація моделі є процесом, який використовується для створення найоптимальнішої та належно налаштованої моделі нейронної мережі, з урахуванням певних пріоритетних обмежень, з метою забезпечення максимальної міцності, ефективності та надійності моделі. Під час оптимізації моделі зберігається її якість, але одночасно досягається покращення швидкодії завдяки використанню попередньо оптимізованих ядер та функцій. Це дозволяє значно прискорити процес виведення висновків з мережі CNN та забезпечити більш ефективне використання ресурсів.

OpenVINO є безкоштовним набором інструментів, який сприяє оптимізації моделей глибокого навчання з використанням інференційного двигуна Intel та їх розгортанню на пристроях Intel. Основні характеристики OpenVINO включають:

– *Оптимізація моделей.* OpenVINO дозволяє оптимізувати моделі глибокого навчання з різних фреймворків, таких як TensorFlow, Caffe, PyTorch та інші, для ефективного використання на пристроях Intel.

– *Інференційний двигун.* OpenVINO надає інференційний двигун, який дозволяє виконувати інференс моделей глибокого навчання на пристроях Intel з високою продуктивністю.

– *Кроссплатформеність.* OpenVINO підтримує різні операційні системи, включаючи Windows, Linux і macOS, що дозволяє розгортати моделі глибокого навчання на різних платформах.

– *Спрощена установка і управління залежностями.* OpenVINO розроблений з мінімальною кількістю залежностей, що спрощує процес установки та управління залежностями.

– *Підтримка різних апаратних платформ.* OpenVINO підтримує різні апаратні платформи Intel, включаючи процесори, графічні прискорювачі, FPGA та VPU, що дозволяє використовувати оптимізовані моделі глибокого навчання на різних пристроях [39].

OpenVINO [40] надає кілька інструментів, які були розроблені з метою допомогти моделям нейромереж працювати швидше та використовувати менше пам'яті. Для оптимізації моделі обрано наступні інструменти: оптимізатор моделі, інструмент оптимізації після навчання (POT).

Оптимізатор моделі виконує конвертацію моделей з різних фреймворків у формат OpenVINO або Intermediate Representation (IR), що дозволяє моделі працювати ефективніше на обладнанні Intel. Цей процес оптимізує топологію мережі з метою поліпшення продуктивності, ефективного використання простору та зменшення обчислювальних вимог до конкретного прискорювача. Важливо зазначити, що оптимізатор моделі не впливає на точність моделі.

Інструмент POT дозволяє прискорити швидкість виведення IR-моделі шляхом застосування квантування після навчання. Це допомагає моделі працювати швидше та використовувати менше пам’яті.

Квантування в нейронних мережах є процесом зміни формату вагових коефіцієнтів і значень активації з високоточного формату (наприклад, 32-розрядні числа з плаваючою комою) на менш точний формат (наприклад, 8-розрядні цілі числа). Квантування дозволяє нейронним мережам працювати вдвічі швидше або навіть у чотири рази швидше, ніж мережам, оптимізованим за допомогою Model Optimizer. Крім того, це зменшує простір, необхідний для збереження мережі на диску, що є ще однією перевагою квантування [40]. На рисунку 3.6 зображено діаграму випадку використання інструменту POT.

Робочий процес OpenVINO:

– *Навчання.* Процес у якому нейромережева модель навчається за допомогою коду;

– *Оптимізатор моделі.* Процес, в якому модель передається в оптимізатор, який має на меті оптимізувати модель та створити проміжне представлення моделі. Під час оптимізації моделей використовуються різні методи, такі як квантування, заморожування, злиття та інші. На цьому етапі попередньо навчені моделі налаштовуються відповідно до обраного фреймворку і перетворюються за допомогою простої однорядкової команди. Користувачі можуть вибирати з широкого спектру попередньо підготовлених моделей, які доступні у OpenVINO Model Zoo. Цей набір моделей містить моделі для різних цілей, включаючи виявлення об'єктів, розпізнавання тексту та оцінку пози людини;

– *Механізм логічного висновку.* Даний механізм працює з проміжним представленням, яке надходить до нього. Основне завдання механізму логічного висновку полягає в перевірці сумісності моделі з використовуваною для її навчання інфраструктурою та апаратним забезпеченням. OpenVINO підтримує такі фреймворки, як TensorFlow, TensorFlow Lite, Caffe, MXNet, ONNX і Kald;

– *Розгортання.* Додаток розгортається на пристроях шляхом забезпечення повного керування пристроями для автоматизованого та надійного розгортання в масштабі [41].

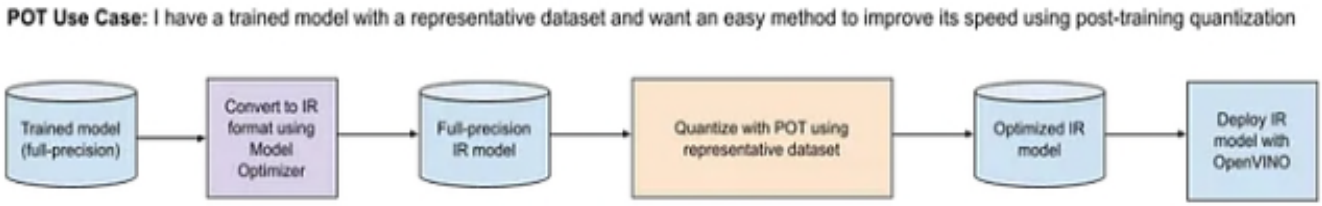


Рисунок 3.6 – Діаграма використання інструменту POT [40]

3.4 Прикладне тестування додатку запропонованого методу

У розробці будь-якого програмного продукту важливим етапом є тестування функціоналу системи. Цей етап є обов'язковим і допомагає розробникам отримати інформацію про коректну роботу програмного додатку відповідно до встановлених вимог. Для тестування функціональності веб-додатку методу виявлення зовнішніх проявів насильства у відеопотоці нейромережевими засобами створено тест-кейси, з метою перевірки базового функціоналу.

Таблиця 3.1 – Тест-кейс 01

Тест-кейс №: 1	Пріоритет: 1	Створено: 01.11.2023, Муляр Е. Р.
Назва: Перевірка запуску веб-додатка		
Кроки		Очікуваний результат
1. Запустити веб-додаток 2. Перейти за посиланням на головну сторінку 3. Перевірити результат		Головна сторінка веб-додатку завантажена
Результат проходження тест-кейсу: успішний		

Таблиця 3.2 – Тест-кейс 02

Тест-кейс №: 2	Пріоритет: 1	Створено: 01.11.2023, Муляр Е. Р.
Назва: Перевірка завантаження відеоматеріалу		
Кроки		Очікуваний результат

1. Запустити веб-додаток 2. Перейти за посиланням на головну сторінку 3. Натиснути на кнопку «CHOOSE VIDEO FILE» 4. Обрати відеоматеріал (файл) з розширенням «.mp4» 5. Перевірити результат	Завантажується відеоматеріал, після чого він відтворюється та відображається у «вікні» додатку
Результат проходження тест-кейсу: успішний	

Таблиця 3.3 – Тест-кейс 03

Тест-кейс №: 3	Пріоритет: 1	Створено: 01.11.2023, Муляр Е. Р.
Назва: Перевірка завантаження прямої відеотрансляції		
Кроки		Очікуваний результат
1. Запустити веб-додаток 2. Перейти за посиланням на головну сторінку 3. Ввести у поле «Enter URL address» посилання прямої відеотрансляції 4. Натиснути на кнопку відправки (знаходиться біля поля вводу) 5. Перевірити результат		Завантажується пряма відеотрансляція, після чого вона відтворюється та відображається у «вікні» додатку
Результат проходження тест-кейсу: успішний		

Таблиця 3.4 – Тест-кейс 04

Тест-кейс №: 4	Пріоритет: 1	Створено: 01.11.2023, Муляр Е. Р.
Назва: Перевірка роботи запропонованого методу при позитивному сценарію (відеоматеріал без насильства)		
Кроки		Очікуваний результат
1. Запустити веб-додаток 2. Перейти за посиланням на головну сторінку		Відтворюється завантажений відеоматеріал, зверху в лівому кутку відображається зелений

3. Натиснути на кнопку « CHOOSE VIDEO FILE » 4. Обрати відеоматеріал (файл) з розширенням «.mp4» та контентом без насильства 5. Перевірити результати	прямокутник з надписом «Peaceful» та значенням менше рівне 0.5.
Результат проходження тест-кейсу: успішний	

Таблиця 3.5 – Тест-кейс 05

Тест-кейс №: 5	Пріоритет: 1	Створено: 01.11.2023, Муляр Е. Р.
Назва: Перевірка роботи запропонованого методу при негативному сценарію (відеоматеріал без ознак насильства)		
Кроки		Очікуваний результат
1. Запустити веб-додаток 2. Перейти за посиланням на головну сторінку 3. Натиснути на кнопку « CHOOSE VIDEO FILE » 4. Обрати відеоматеріал (файл) з розширенням «.mp4» та контентом з явними ознаками насильства 5. Перевірити результати		Відтворюється завантажений відеоматеріал, зверху в лівому кутку відображається червоний прямокутник з надписом «Suspicious» та значенням більше 0.5.
Результат проходження тест-кейсу: успішний		

Висновки до розділу 3

В третьому розділі описано реалізацію веб-додатку методу виявлення зовнішніх проявів насильства нейромережевими засобами. Описано компоненти ІС, а також їх функціонал. Продемонстровано реалізацію цих компонентів у вигляді програмного коду. Наведено бібліотеки та інформаційні технології, які використовувались для реалізації додатку та відповідно методу.

Проведено прикладне тестування інформаційної системи, що призвело до створення набору тест-кейсів для перевірки функціональності інтерфейсу, з яким взаємодіє користувач, а також функціональної частини веб-додатку. Крім того, описано окремі тест-кейси для перевірки роботи самого методу. В результаті всі компоненти успішно визнані справними.

РОЗДІЛ 4

Дослідження методу виявлення зовнішніх проявів насильства у відеопотоці нейромережевими засобами

4.1 Набір даних

Важливим та остаточним етапом у реалізації запропонованого методу є визначення його ефективності роботи. Для того, щоб оцінити ефективність роботи запропонованого методу, проведено декілька експериментів щодо трьох наборів даних: «Hockey fights» [42], «Violent flows» [43], «Livestream» [44]. Кожен набір даних розділений на 4 рівних частини: 3 частини використовуються для навчання моделі, а одна зарезервована для тестування. Оцінка ефективності методу базується на визначенні загальної точності для кожного набору даних. У таблиці 4 наведено детальний опис використаних наборів даних для проведення експериментів.

Таблиця 4 – Детальний опис використаних наборів даних

Datasets	Samples	Resolution	Violent Scenes		Non-Violent Scenes	
			No. of Clips	Frame Rate	No. of Clips	Frame Rate
Hockey fights	1000	360x288	500	30	500	30
Violent flows	246	360x240	146	30	100	30
Livestream	80	640x480	50	30	30	30

Набір даних «Hockey fights» був створений Nievas і включає в себе 1000 коротких відеокліпів, отриманих із Національної хокейної ліги (НХЛ). У даному наборі даних 500 відеокліпів позначено як випадки бійок, і 500 відеокліпів позначено як випадки, коли бійок немає. Кожен відеоролик містить 30 кадрів із роздільною здатністю 360×288 пікселів. У класі з бійками всі відеоролики пов'язані із сценами бійок на хокейних майданчиках, а клас без бійки також включає в себе відеокліпи, що

не містять бійок, але знімалися в тому ж самому спортивному середовищі для надійного виявлення сцен насильства у спортивних відео [42].

Набір даних «Violent Flows» представляє собою колекцію реальних відеозаписів, що відображають насильство в натовпі, а також включає стандартні протоколи тестування для оцінки класифікації подій як насильства або не насильства та виявлення насильницьких випадків. Цей набір даних містить 246 відеороликів, які були отримані з YouTube. Найкоротший ролик має тривалість 1,04 секунди, найдовший – 6,52 секунди, і середня тривалість відеоролика становить 3,60 секунди [43].

Набір даних "Livestream" складається з прямих відеотрансляцій з платформи Twitch. Всього в цьому наборі даних міститься 80 відеотрансляцій, серед яких 50 включають в себе сцени бійок, а 30 інших трансляцій не мають подібних випадків насильства.

4.2 Визначення загальної точності методу виявлення зовнішніх проявів насильства у відеопотоці

Для визначення загальної точності роботи методу на відповідному набору даних застосовано наступний підхід:

Етап 1 – Визначення середньої точності

Необхідно зі всіх точностей (accuracy) знайти точність з найбільшим значенням, знаючи цю точність та її епоху потрібно сформувати новий масив даних точностей, які знаходяться в радіусі 10 епох від цієї максимальної точності. Отримавши масив точностей можна знайти середню точність за допомогою наступної формули:

$$A = \frac{\sum_{i=1}^{N} x_i}{N}, \tag{22}$$

де A – середня точність, N – загальна кількість влучень, x – значення відповідної точності, i – порядковий номер.

Етап 2 – Визначення стандартного відхилення

Отримавши середню точність можна знайти стандартне відхилення за наступною формулою:

$$\sigma = \sqrt{\frac{\sum_{i=1}^{N}(x_i - A)^2}{N}}, \tag{23}$$

де σ - стандартне відхилення, N – загальна кількість влучень, x – значення відповідної точності, i – порядковий номер, A – середня точність.

Таким чином, отримане значення середньої точності буде відповідати загальній точності, значення стандартного відхилення буде відповідати похибці середнього значення, яку можна виразити як ± значення, тобто загальна точність = середня точність ± стандартне відхилення.

Загальна точність для набору даних «Hockey fights» склала 98.5±0.78%, зображено на рисунку 4.1. У таблиці 4.1 наведено порівняння точності методів для набору даних «Hockey fights».

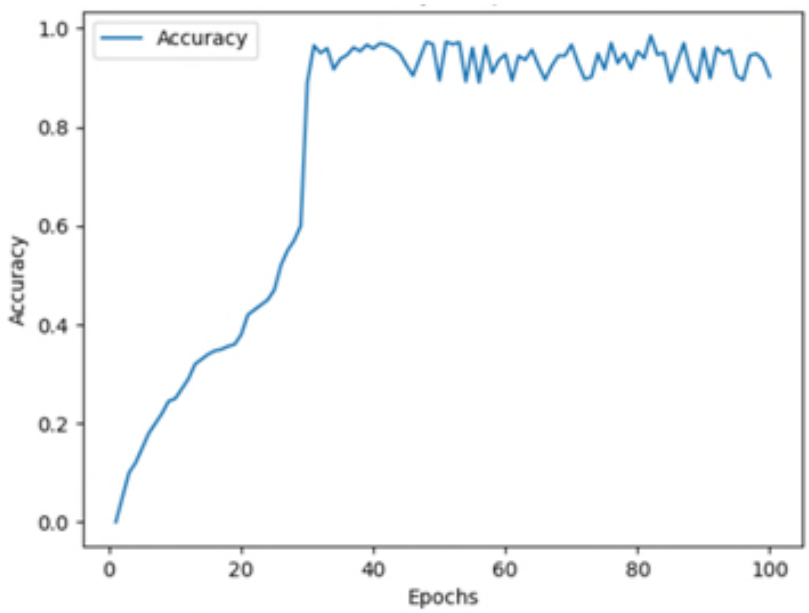


Рисунок 4.1 – Загальна точність набору даних «Hockey fights»

Таблиця 4.1 – Порівняння точності методів на наборі даних «Hockey fights»

Метод	Hockey fights
Запропонований	98.5 ± 0.78%
Просторово-часовий кодер [20]	96.54 ± 1.01%
3D CNN [22]	98.3 ± 0.81%

Загальна точність для набору даних «Violent flows» склала 99.45±0.37%, зображено на рисунку 4.2. Наведено у таблиці 4.2 порівняння точності методів для набору даних «Violent flows».

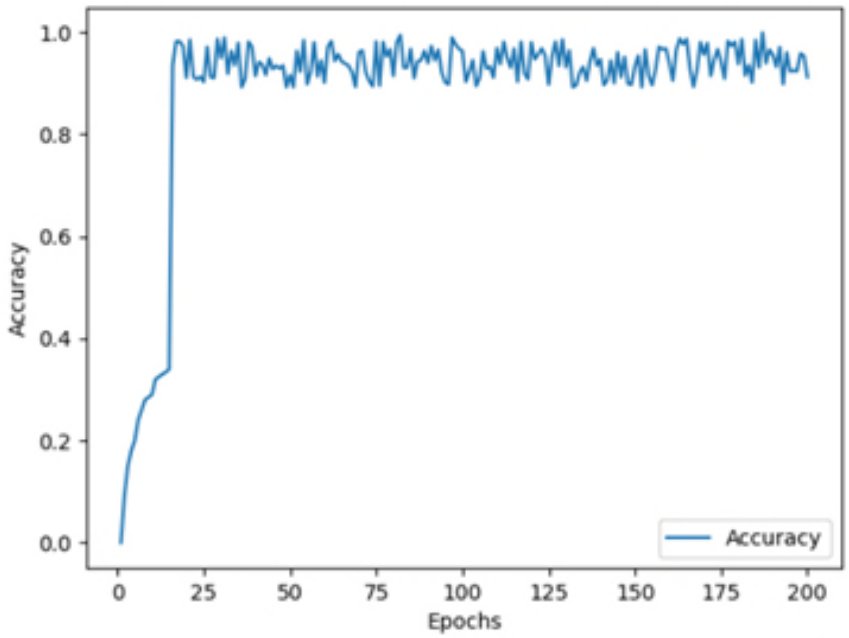


Рисунок 4.2 – Загальна точність набору даних «Violent flows»

Таблиця 4.2 – Порівняння точності методів на наборі даних «Violent flows»

Метод	Violent flows
Запропонований	99.45 ± 0.37%
Просторово-часовий кодер [20]	92.18 ± 3.29%

3D CNN [22]	97.17 ± 0.95%

Загальна точність для набору даних «Livestream» склала 87.4±2.19%, зображено на рисунку 4.3. У таблиці 4.3 зображено порівняння точності методів для набора даних «Livestream».

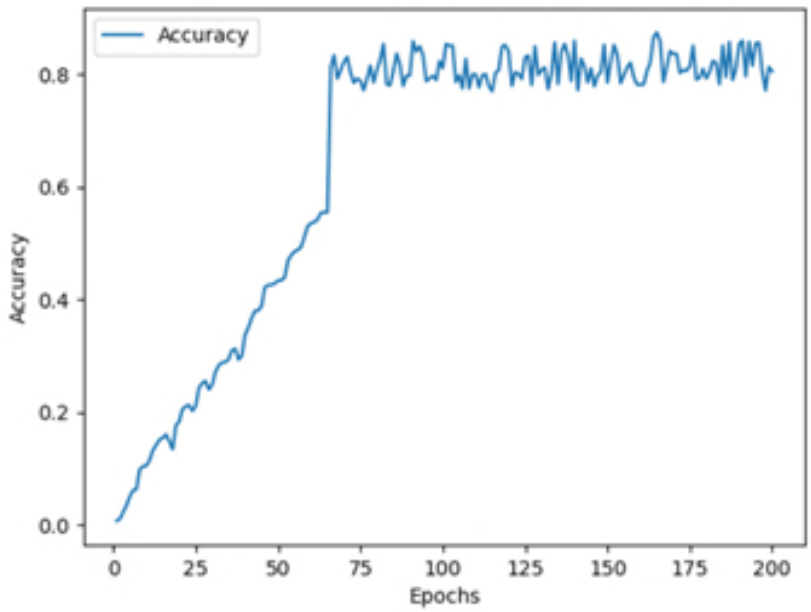


Рисунок 4.3 – Загальна точність набору даних «Livestream»

Таблиця 4.3 – Порівняння точності методів на наборі даних «Livestream»

Метод	Livestream
Запропонований	87.4 ± 2.19%
Просторово-часовий кодер [20]	-
3D CNN [22]	-

Особливістю запропонованого методу є робота з відеоматеріалом (відеопотоком) у реальному часі. Дана можливість досягається за рахунок того, що згорткова нейронна мережа навчена на неперервному потоку даних з мультимедійних платформ для онлайн трансляцій використовуючи метод fine-tuning. Тобто, навчання

відбувається в режимі реального часу і триватиме, доки примусово не зупиниться трансляція. Відповідно, тестування запропонованого методу відбувалося аналогічним чином, методу на вхід подавалася трансляція і він в реальному часі видавав оцінку сцени, яка відображалася на трансляції.

Висновки до розділу 4

В четвертому розділі виконано дослідження методу виявлення зовнішніх проявів насильства у відеопотоці нейромережевими засобами. Під час виконання зібрано три набори даних: «Hockey fights» та «Violent flows» (в якості статичного відеоматеріалу), «Livestream» (в якості динамічного відеоматеріалу), які використовувалися для навчання та тестування запропонованого методу. Наведено характеристичні параметри (розмір, кількість тощо) зібраних датасетів.

Проведено декілька експериментів з метою визначення загальної точності роботи запропонованого методу на відповідних наборах даних. Для набору даних «Hockey fights» загальна точність становить 98.5±0.78%, для «Violent flows» склала 99.45±0.37%, для «Livestream» відповідно 87.4±2.19%. Результати наведені у вигляді графік.

Порівняно метод виявлення зовнішніх проявів насильства у відеопотоці нейромережевими засобами з іншими підходами. Наведено таблиці у яких вказані загальні точності підходів на відповідних наборах даних.

Загальні висновки

В результаті кваліфікаційної роботи магістра розроблено метод виявлення зовнішніх проявів насильства у відеопотоці нейромережевими засобами.

В процесі розробки методу виконано наступні задачі:

– Досліджено предметну область поставленого завдання та проаналізовано існуючі підходи та публікації з метою доведення актуальності проблеми, яку необхідно вирішити.

– Розроблено метод виявлення зовнішніх проявів насильства у відеопотоці нейромережевими засобами. Запропонований метод складається з 2 частин: згорткова нейронна мережа (задача полягає у вилученні та формуванні набору ознак) та SVM (задача полягає у розрахуванні оцінки щодо ймовірності належності вхідних даних (ознак) до певного класу (насильницького або не насильницького)). Для того, щоб метод міг працювати як зі статичним відеоматеріалом (звичайне відео) так і з динамічним (пряма відеотрансляція) обрано наступні методи навчання: метод зворотного поширення помилки та метод тонкого налаштування (fine-tuning).

– Розроблено веб-додаток запропонованого методу. Для frontend частини обрано фреймворки Electron.js та Bootstrap 4, для backend частини обрано фреймворк Django, для реалізації методу обрано мову програмування Python та відповідні бібліотеки для роботи з нейромережевими засобами.

Проведено валідацію розробленого методу на тестових наборах даних: «Hockey fights» та «Violent flows» (в якості статичного відеоматеріалу), «Livestream» (в якості динамічного відеоматеріалу). Для набору даних «Hockey fights» загальна точність становить 98.5±0.78%, для «Violent flows» склала 99.45±0.37%, для «Livestream» відповідно 87.4±2.19%.

Перелік посилань

1. Згорткові нейронні мережі (CNN) [Електронний ресурс]. – Режим доступу: https://techukraine.net/згорткові-нейронні-мережі-cnn-вступ/
2. Violence a global public health problem [Електронний ресурс]. – Режим доступу: https://www.scielo.br/j/csc/a/3hrn64cpBqBFb9mNfP4KGXr/?lang=en
3. Форми насильства [Електронний ресурс]. – Режим доступу: https://1547.ukc.gov.ua/knowdledge/yaki-formy-domashnogo-nasylstva/
4. Форми насильства: фізичне, психологічне, економічне [Електронний ресурс]. – Режим доступу: https://www.hsa.org.ua/blog/formy-nasylstva-fizychne-psyhologichne-ekonomichne-yak-dovesty-fakt-kozhnogo-z-perelichenyh-riznovydiv-nasylstva
5. Weakly Supervised Violence Detection in Surveillance Video [Електронний ресурс]. – Режим доступу: https://www.mdpi.com/1424-8220/22/12/4502
6. Efficient Violence Detection in Surveillance [Електронний ресурс]. – Режим доступу: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC8950857/
7. Штучна нейронна мережа [Електронний ресурс]. – Режим доступу: https://uk.wikipedia.org/wiki/Штучна_нейронна_мережа
8. Нейромережеві технології у відеоспостереженні [Електронний ресурс]. – Режим доступу: https://www.dssl.ru/publications/obzory/epokha-neyrosetevykh-tekhnologiy-v-videonablyudenii-nachalas/
9. Action Recognition from Videos using Deep Neural Networks [Електронний ресурс]. – Режим доступу: https://escholarship.org/uc/item/2mr798mn
10. ЗГОРТКОВІ НЕЙРОННІ МЕРЕЖІ, ВИЯВЛЕННЯ ОБ’ЄКТІВ, ГЛИБИННІНЕЙРОННІ МЕРЕЖІ, МАШИННЕ НАВЧАННЯ, ЛИЦЬОВІ МАСКИ, ФУНКЦІЯ ВТРАТ, АЛГОРИТМ ОПТИМІЗАЦІЇ [Електронний ресурс]. – Режим доступу: https://openarchive.nure.ua/server/api/core/bitstreams/abf36e40-7e62-40b4-87b0-8cc130a64074/content

11. Оптимізація та побудова системи генерувань зображень на основі генеративно-змагальних мереж [Електронний ресурс]. – Режим доступу: https://ai.kpi.ua/ua/masters/thesis/28521smai-dyshkant_magistr.pdf
12. Штучна нейронна система з використанням генетичних алгоритмів для її тренування [Електронний ресурс]. – Режим доступу: https://dspace.nau.edu.ua/bitstream/NAU/56999/1/ФККПІ_2021_211_Саранча_Р_М.pdf
13. Застосування методів реідентифікації осіб для аналізу матеріалів систем відеоспостереження [Електронний ресурс]. – Режим доступу: https://ela.kpi.ua/bitstream/123456789/31293/1/Merzlikina_bakalavr.pdf
14. Action Recognition from Videos using Deep Neural Networks [Електронний ресурс]. – Режим доступу: https://escholarship.org/uc/item/2mr798mn
15. Рекурентна нейронна мережа (RNN): види, навчання [Електронний ресурс]. – Режим доступу: https://neurohive.io/ru/osnovy-data-science/rekurrentnye-nejronnye-seti/
16. Згорткова нейронна мережа [Електронний ресурс]. – Режим доступу: https://uk.wikipedia.org/wiki/Згорткова_нейронна_мережа
17. Згорткові нейронні мережі (CNN) [Електронний ресурс]. – Режим доступу: https://techukraine.net/згорткові-нейронні-мережі-cnn-вступ/
18. Simonyan, K., Zisserman, A.: Very deep convolutional networks for large-scale image recognition [Електронний ресурс]. – Режим доступу: http://arxiv.org/abs/1409.1556
19. Complete Guide To Bidirectional LSTM [Електронний ресурс]. – Режим доступу: https://analyticsindiamag.com/complete-guide-to-bidirectional-lstm-with-python-codes/
20. Bidirectional Convolutional LSTM for the Detection of Violence in Videos [Електронний ресурс]. – Режим доступу: https://openaccess.thecvf.com/content_ECCVW_2018/papers/11130/Hanson_Bidire

ctional_Convolutional_LSTM_for_the_Detection_of_Violence_in_Videos_ECCVW_2018_paper.pdf

21. 3D Convolutions: Understanding + Use Case [Електронний ресурс]. – Режим доступу: https://www.kaggle.com/code/shivamb/3d-convolutions-understanding-use-case
22. Efficient Violence Detection Using 3D Convolutional Neural Networks [Електронний ресурс]. – Режим доступу: https://www.researchgate.net/profile/Tanfeng-Sun/publication/337537845_Efficient_Violence_Detection_Using_3D_Convolutional_Neural_Networks/links/5f75e252a6fdcc00864ccb95/Efficient-Violence-Detection-Using-3D-Convolutional-Neural-Networks.pdf
23. Convolutional Neural Networks from the ground up [Електронний ресурс]. – Режим доступу: https://towardsdatascience.com/convolutional-neural-networks-from-the-ground-up-c67bb41454e1
24. SVM. Детальний розбір методу опорних векторів, реалізація на Python [Електронний ресурс]. – Режим доступу: https://habr.com/ru/companies/ods/articles/484148/
25. Derivation of Backpropagation in Convolutional Neural Network (CNN) [Електронний ресурс]. – Режим доступу: https://www.pycodemates.com/2023/07/backward-pass-in-convolutional-neural-network-explained.html
26. Веб-платформа [Електронний ресурс]. – Режим доступу: http://x-site.by/products/web-platform
27. Electron [Електронний ресурс]. – Режим доступу: https://uk.wikipedia.org/wiki/Electron
28. Bootstrap та його основні компоненти [Електронний ресурс]. – Режим доступу: http://intech.lviv.ua/bootstrap-та-його-основні-компоненти/
29. Django [Електронний ресурс]. – Режим доступу: https://uk.wikipedia.org/wiki/Django

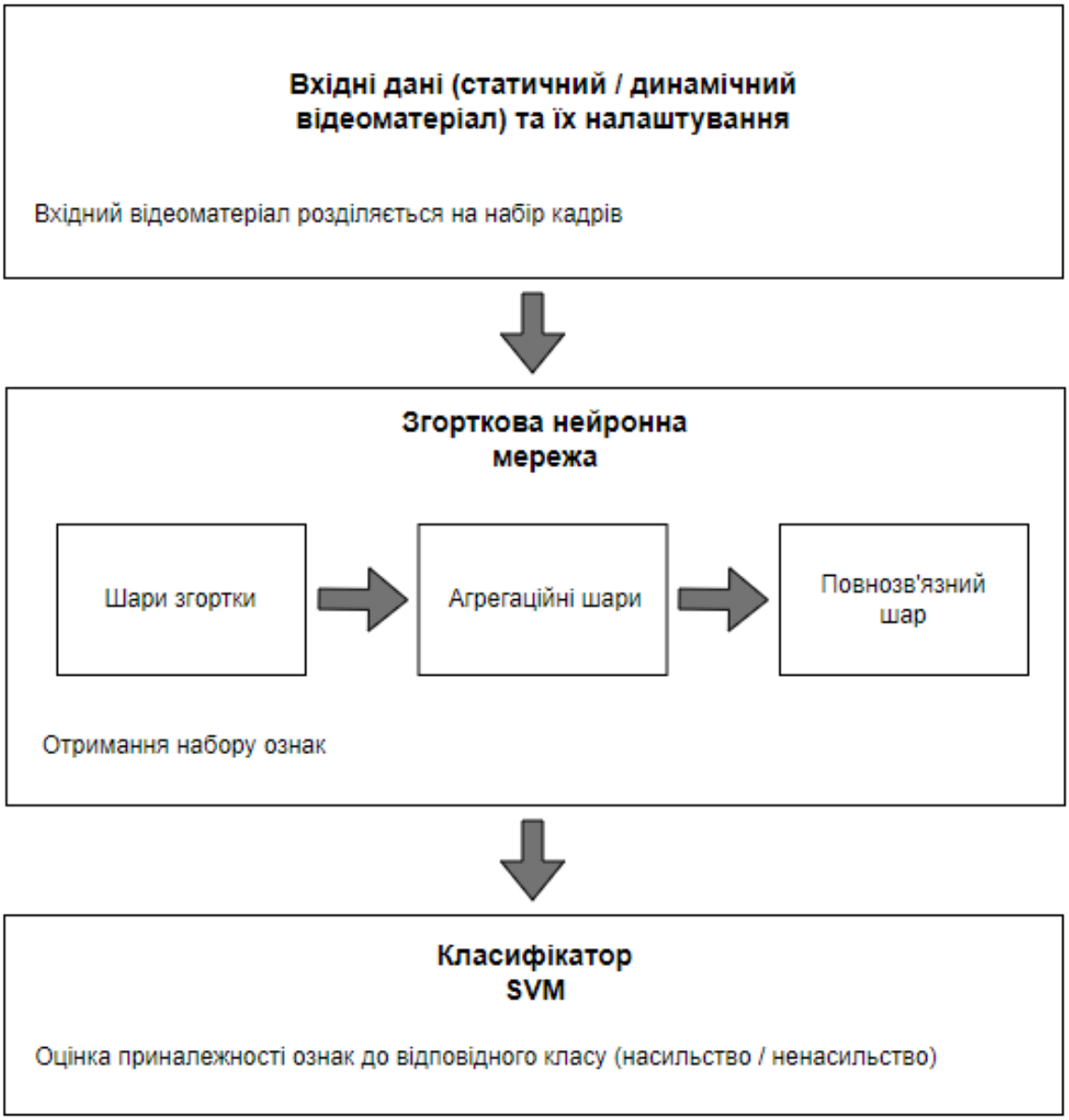
30. Особливості та плюси використання Django [Електронний ресурс]. – Режим доступу: https://habr.com/ru/sandbox/156526/
31. Python [Електронний ресурс]. – Режим доступу: https://uk.wikipedia.org/wiki/Python
32. ЩО ТАКЕ МОВА ПРОГРАМУВАННЯ PYTHON? [Електронний ресурс]. – Режим доступу: https://freehost.com.ua/ukr/faq/wiki/chto-takoe-jazik-programmirovanija-python/
33. TensorFlow [Електронний ресурс]. – Режим доступу: https://uk.wikipedia.org/wiki/TensorFlow
34. Keras [Електронний ресурс]. – Режим доступу: https://keras.io
35. OpenCV [Електронний ресурс]. – Режим доступу: https://en.wikipedia.org/wiki/OpenCV
36. NumPy [Електронний ресурс]. – Режим доступу: https://en.wikipedia.org/wiki/NumPy
37. Scikit-learn [Електронний ресурс]. – Режим доступу: https://en.wikipedia.org/wiki/Scikit-learn
38. Matplotlib [Електронний ресурс]. – Режим доступу: https://en.wikipedia.org/wiki/Matplotlib
39. Що таке OpenVINO? [Електронний ресурс]. – Режим доступу: https://habr.com/ru/companies/intel/articles/546438/
40. Developer Guide: Model Optimization with the OpenVINO™ Toolkit [Електронний ресурс]. – Режим доступу: https://medium.com/openvino-toolkit/developer-guide-model-optimization-with-the-openvino-toolkit-d19a201dd3ce
41. What is OpenVINO? – The Ultimate Overview in 2024 [Електронний ресурс]. – Режим доступу: https://viso.ai/computer-vision/intel-openvino-toolkit-overview/
42. Hockey Fight Detection Dataset [Електронний ресурс]. – Режим доступу: https://paperswithcode.com/dataset/hockey-fight-detection-dataset

43. Violent-Flows Dataset [Електронний ресурс]. – Режим доступу: https://paperswithcode.com/dataset/violent-flows
44. Livestream Dataset [Електронний ресурс]. – Режим доступу: https://www.twitch.tv/directory/category/just-chatting

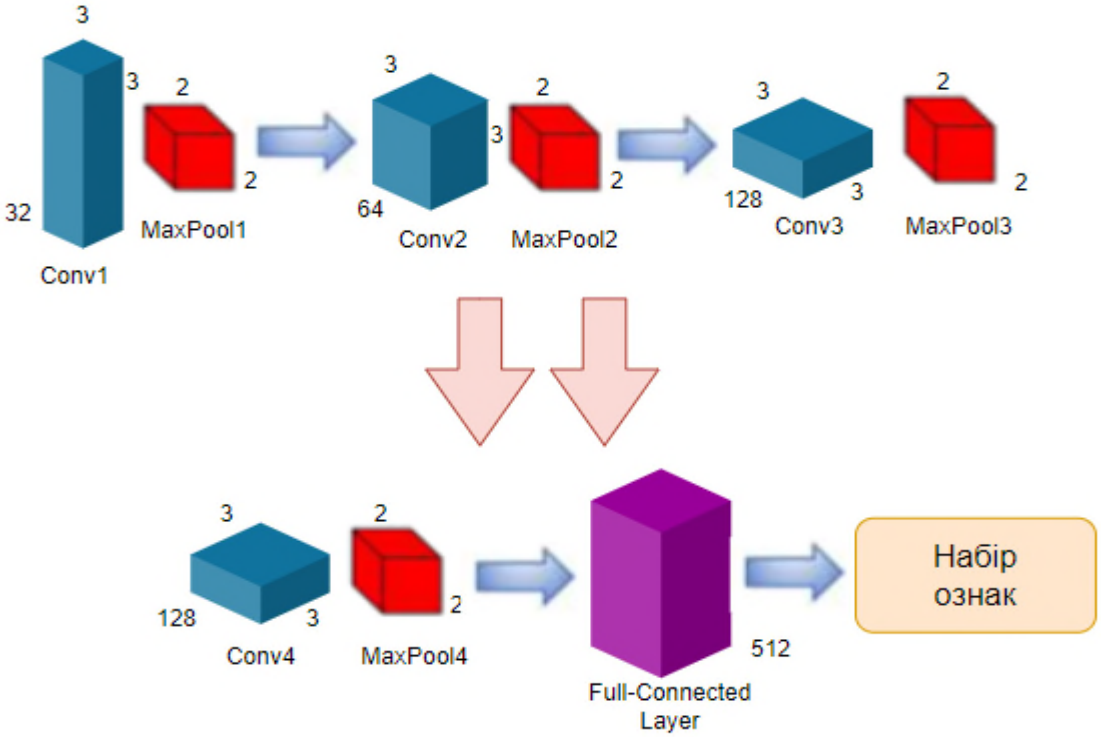
ДОДАТКИ

Додаток А

Схема методу виявлення зовнішніх проявів насильства у відеопотоці нейромережевими засобами



Архітектура згорткової нейронної мережі в контексті запропонованого методу



Додаток Б

Світлини наукових публікацій, виконаних при роботі над кваліфікаційною роботою магістра

(ксерокопії титульної сторінки, сторінки змісту та всіх сторінок із публікацією)

Перелік наукових публікацій:

Збірник наукових праць – Метод виявлення ознак насильства у відеоматеріалах нейромережевими засобами / Муляр Е.Р., Багрій Р.О., Манзюк Е.А., Пасічник О.А. // Збірник наукових праць за матеріалами Всеукраїнської науково-практичної конференції «Актуальні проблеми комп'ютерних наук – 2023» Хмельницький, 2023.

Науковий журнал – Метод виявлення зовнішніх проявів насильства у відеопотоці нейромережевими засобами / Муляр Е.Р., Багрій Р.О., Манзюк Е.А., Пасічник О.А. // Науковий журнал «Вісник Хмельницького національного університету. Серія: Технічні науки» Хмельницький, 2023.

Хмельницький національний університет

XV ВСЕУКРАЇНСЬКА НАУКОВО-ПРАКТИЧНА КОНФЕРЕНЦІЯ

АКТУАЛЬНІ ПРОБЛЕМИ КОМП'ЮТЕРНИХ НАУК
АПКН–2023

17-18 листопада 2023 р.

м. Хмельницький

УДК 004.8

Муляр Е.Р., Багрій Р.О., Пасічник О.А., Манзюк Е.А.

Хмельницький національний університет

МЕТОД ВИЯВЛЕННЯ ОЗНАК НАСИЛЬСТВА У ВІДЕОМАТЕРІАЛАХ НЕЙРОМЕРЕЖЕВИМИ ЗАСОБАМИ

Проблема виявлення насильства у відеопотоці є актуальною в сучасному світі, де зростає кількість відеоматеріалів з насильницькими сценами. Виявлення та реагування на такі сцени мають велике значення для забезпечення безпеки громадських просторів та захисту прав людини. Запропоновано метод виявлення ознак насильства у відеоматеріалах нейромережевими засобами. Метод відкриває шлях до розв'язання завдань виявлення ознак насильства на відеоматеріалах у реальному часі, а також фільтрації контенту на потокових мультимедійних платформах.

The problem of detecting violence in a video stream is relevant in the modern world, where the number of videos with violent scenes is growing. Detecting and responding to such scenes is important for ensuring the safety of public spaces and protecting human rights. The paper proposes a method for detecting signs of violence in video materials using neural network tools. The method opens the way to solving the problems of detecting signs of violence in real-time video materials, as well as filtering content on streaming multimedia platforms.

Вступ

Сьогодні для протидії такій суспільній проблемі як насильство розпочали активно використовувати системи відеоспостережень. Такі країни, як Китай та Південна Корея, є лідерами у встановленні камер відеоспостереження, і результати використання цих систем є вражаючими. У Китаї рівень насильства у громадських місцях знизився на 60%, а в Південній Кореї - на приблизно 50% [1]. Однак, ці системи мають деякі недоліки. Основною проблемою є людський фактор, зокрема неуважність та недбалість операторів-спостерігачів. Зазвичай оператор може ефективно контролювати лише обмежену кількість камер відеоспостереження. Однак, коли кількість камер перевищує їх межі, оператори можуть допускати помилки або пропускати випадки насильства. Для вирішення цієї проблеми сьогодні активно використовуються передові інформаційні технології, зокрема штучні нейронні мережі. Завдяки цим технологіям, системи відеоспостереження можуть автоматично аналізувати великий обсяг відеоданих і виявляти потенційні випадки насильства. Штучний

інтелект допомагає зменшити навантаження на операторів і забезпечує більш точне та ефективне виявлення подій [2].

Аналіз існуючих публікацій

Для реалізації інтелектуального відеоспостереження сьогодні активно використовують підходи модифікації та поєднання різних методів та архітектур нейромереж [3].

Прикладом такого підходу є модель нейронних мереж BiConvLSTM (Bidirectional Convolutional LSTM). Ця модель поєднує в собі два потужних компоненти: двосторонню згорткову мережу (BiConv) та рекурентну нейронну мережу LSTM (Long Short-Term Memory). BiConvLSTM використовується для аналізу послідовних даних, таких як зображення або відео, з метою виявлення шаблонів та залежностей у цих даних [4].

Іншим прикладом використання нейронних мереж для задачі виявлення насильства у відеопотоці є 3D CNN (3D Convolutional Neural Network). Дана модель являється згортковою нейронною мережею, що використовується для обробки тривимірних даних, таких як відео, медичні зображення або тривимірні моделі [5].

Приклад використання моделі BiConvLSTM наведено у роботі [6], де розглянуто підхід до виявлення насильства у відео за допомогою методу «просторово-часовий кодер». «Просторово-часовий кодер» представляє собою метод який побудований на основі декількох архітектур: BiConvLSTM, VGG13 [7] для кодування кожного відеокадру як набір карт функцій. Ці карти функцій потім передаються до BiConvLSTM для подальшого кодування у часовому напрямку відео. Виконується поелементна максимізація кодувань для створення представлення відео, яке передається класифікатору для виявлення насильства. Щодо результатів роботи даного підходу, то для набору даних «Hockey fights» точність склала 96.54%, для набору даних «Violent flows» - 92.18%.

У роботі [8] запропоновано метод до виявлення насильства у відео за допомогою 3D-CNN. 3D-CNN спочатку обробляє кожний кадр відео, використовуючи набори фільтрів для виявлення важливих ознак, таких як рух, форма та колір. Потім 3D-CNN обробляє послідовність кадрів, використовуючи 3D-фільтри. Класифікація здійснюється шляхом застосування логістичної регресії до вихідного тензора 3D-CNN. На рахунок результатів роботи даного методу, то для набору даних «Hockey fights» точність склала 98.3%, для набору даних «Violent flows» - 97.17%.

В наведених публікаціях не розглянуто підходи до взаємодії запропонованих методів з відеоматеріалами у реальному часі, тому виникають питання в ефективності даних підходів при роботі з відеопотоком даних в реальному часі.

Метою роботи є розробка методу для виявлення ознак насильства у відеоматеріалах нейромережевими засобами. Метод повинен працювати як зі статичним відеоматеріалом (відеоролик) так і з динамічним (відеопотік в реальному часі).

Метод виявлення ознак насильства у відеоматеріалах нейромережевими засобами

Робота запропонованого методу полягає в отриманні ознак насильства з кадрів вхідного відео за допомогою згорткової нейронної мережі і визначення ступеню насильства у відсотковому відношенні у відеопотоці за допомогою SVM (методу опорних векторів). На рисунку 1 представлено архітектуру даного методу.

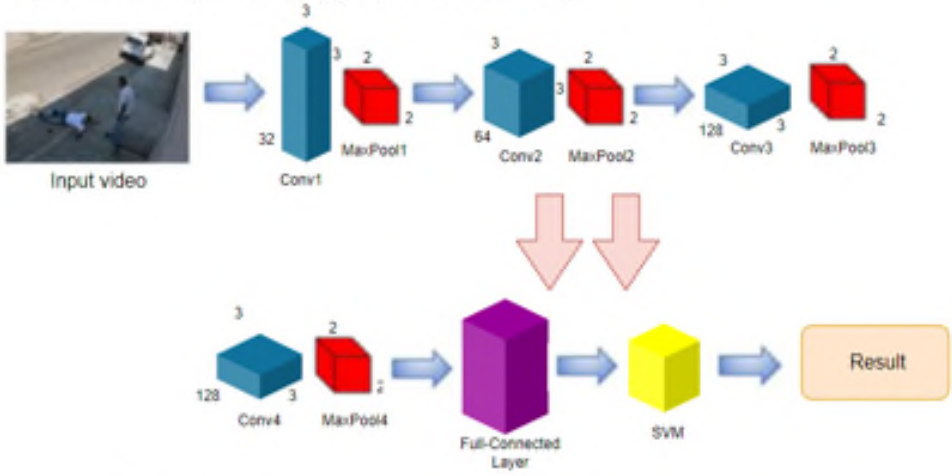


Рисунок 1 – Архітектура заборонованого методу

Етап 1 – Отримання кадрів із вхідного відеоматеріалу

Необхідно розбити вхідний відеоматеріал на послідовність кадрів та перетворити кожен кадр у карту зображень.

Етап 2 – Операція згортки

На даному етап необхідно виконати операцію згортки на вхідне зображення для того, щоб отримати карту ознак. Для виконання даної операції використовуються фільтри (матриця параметрів). Для даної нейронної мережі було обрано 4 рівня фільтрів по 32, 64, 128, 128 фільтрів на кожному рівні відповідно. Формування карти ознак можна здійснити за допомогою наступної математичної формули:

$$M(i,j) = (K * X)(i,j) = \sum_m \sum_n K(m,n) X(i-m, j-n), \quad (1)$$

де M – елемент карти ознак з координатами i та j, X – вхідне зображення, K – детектор ознак, (m, n) – розмірності детектора ознак.

Етап 3 – Операція максимального об'єднання

Максимальне об'єднання являється операцією, яка об'єднує елементи в межах фільтра на карті ознак і вибирає найбільший елемент. Тобто, після проходження через шар максимального об'єднання, отримується нова карта ознак, яка містить найбільш помітні ознаки з попередньої карти ознак. Виконати дану операцію можна за допомогою наступної формули:

$$p(i,j) = \max_{i,j}(x(i-m, j-n)), \quad (2)$$

де p(i, j) – значення елементу поточного рівня з координатами i та j, x – вхідні дані з попередніх рівнів, (m, n) – розмірність рецептивного поля.

Етап 4 – Повнозв'язний рівень

Повнозв'язний рівень представляє собою модель багаторівневого персептрона, де всі нейрони з наступного шару з'єднані з нейронами попереднього шару. Цей рівень використовуються на передостанньому етапі роботи мережі для підготовки результатів на виході мережі. На даному рівні виконується обчислення скалярного добутку даних та параметрів з додаванням зсуву.

Етап 5 – Класифікація отриманих ознак за допомогою SVM

Метод опорних векторів (SVM) використовується для знаходження параметрів гіперплощини у великому чи нескінченному вимірному просторі, яка може служити для класифікації. Головна ідея полягає в тому, щоб знайти гіперплощину, яка має найбільшу відстань до найближчих навчальних точок будь-якого класу. На виході класифікатор SVM видає оцінку, яка представляє ймовірність того, що вхідні дані належать до певного класу (насильницького або не насильницького).

Для того, щоб запропонований метод міг працювати з відеоматеріалом в реальному часі під час процесу навчання нейронної мережі використовувався метод fine-tuning. Суть даного методу полягає в тому, що нейромережу спочатку навчають на наборі даних готових відео, а потім поступово додають до набору даних відеопотоки у реальному часі. Це допомагає нейромережі навчитися розпізнавати насильство в умовах шуму, швидкої зміни та різноманітності.

Аналіз ефективності створеного методу

Для того, щоб оцінити ефективність роботи запропонованого методу, проведено декілька експериментів щодо двох наборів даних для виявлення насильства: «Hockey fights» [9], «Livestream» [10]. Оцінка ефективності методу базується на визначенні загальної точності для кожного набору даних.

Загальна точність для набору даних «Hockey fights» склала 98.5%, представлено на рисунку 2.

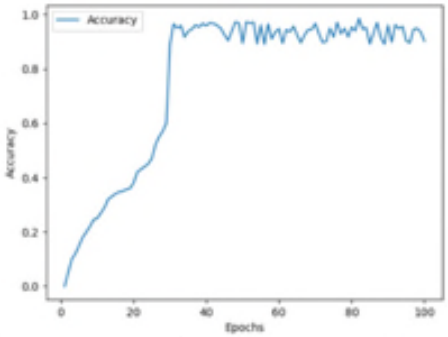


Рисунок 2 – Загальна точність набору даних «Hockey fights»

Загальна точність для набору даних «Livestream» склала 87.4%, представлено на рисунку 3.

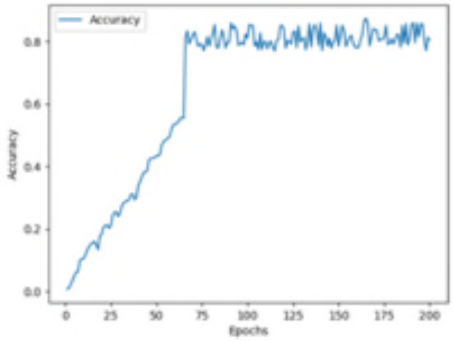


Рисунок 3 – Загальна точність набору даних «Livestream»

Висновки

Отже, запропонований метод для виявлення ознак насильства у відеоматеріалах нейромережевими засобами дозволяє визначити ступінь насильницького характеру у відсотковому відношенні, відповідно на статичних відеоматеріалах так і на динамічних (відеопотік). Подальші дослідження спрямовані на пришвидшення роботи та покращення точності запропонованого методу для набору даних, які пов’язані з насильницькими діями у реальному часі.

Перелік посилань

1. Violence a global public health problem URL: https://www.scielo.br/j/csc/a/3hrn64cpBqBFb9mNfP4KGXr/?lang=en
2. Використання технологій штучного інтелекту протидії злочинності URL: https://ivpz.kh.ua/wp-content/uploads/2020/12/Матеріали-семінару_Використання-техн-штучного-інтел_5.11.2020.pdf
3. A CNN-RNN Combined Structure for Real-World Violence Detection in Surveillance Cameras URL: https://www.mdpi.com/2076-3417/12/3/1021
4. NABNet: A Nested Attention-guided BiConvLSTM network URL: https://www.sciencedirect.com/science/article/abs/pii/S1746809422007017
5. 3D Convolutional Neural Network — A Guide for Engineers URL: https://www.neuralconcept.com/post/3d-convolutional-neural-network-a-guide-for-engineers
6. Convolutional LSTM for the Detection of Violence in Videos URL: https://openaccess.thecvf.com/content_ECCVW_2018/papers/11130/Hanson_Bidirectional_Convolutional_LSTM_for_the_Detection_of_Violence_in_Videos_ECCVW_2018_paper.pdf
7. Simonyan, K., Zisserman, A.: Very deep convolutional networks for large-scale image recognition. In International Conference on Learning Representations (2015) URL: http://arxiv.org/abs/1409.1556
8. Efficient Violence Detection Using 3D Convolutional Neural Networks URL: https://www.researchgate.net/profile/Tanfeng-Sun/publication/337537845_Efficient_Violence_Detection_Using_3D_Convolutional_Neural_Networks/links/5f75e252a6fdcc00864ccb95/Efficient-Violence-Detection-Using-3D-Convolutional-Neural-Networks.pdf
9. Hockey Fight Detection Dataset URL: https://paperswithcode.com/dataset/hockey-fight-detection-dataset
10. Livestream URL: https://www.twitch.tv/directory/category/just-chatting

ISSN 2307-5732
DOI 10.31891/2307-5732

Науковий журнал

ВІСНИК

Хмельницького національного університету

Технічні науки

Довідка: ВХНУ ТН 24/11/23

Видання: Вісник Хмельницького національного університету. Технічні науки

Категорія фаховості видання: фахове видання України, у якому можуть публікуватися результати дисертаційних робіт на здобуття наукових ступенів доктора наук, кандидата наук та ступеня доктора філософії, категорії «Б» філософії, категорії «Б» (наказ МОН №1643 від 28.12.2019, наказ МОН №409 від 17.03.2020).

Напрям – технічні науки за спеціальностями – 101, 121, 122, 123, 124, 125, 141, 151, 161, 172, 181, 182 (28.12.2019), спеціальності – 131, 132, 133 (17.03.2020)

Назва статті: МЕТОД ВИЯВЛЕННЯ ЗОВНІШНІХ ПРОЯВІВ НАСИЛЬСТВА У ВІДЕОПОТОЦІ НЕЙРОМЕРЕЖЕВИМИ ЗАСОБАМИ

Автори: МУЛЯР ЕДУАРД, БАГРІЙ РУСЛАН, ПАСІЧНИК ОЛЕКСАНДР, МАНЗЮК ЕДУАРД (Хмельницький національний університет»)

Номер, у який прийнято статтю: №6 до друку рекомендовано буде до 25 грудня 2023 року.

24.11.2023

Начальник відділу
інтелектуальної власності та трансферу технологій Ю.В.Кравчик

ВІДДІЛ КАДРІВ
І.С.Мартинюк

УДК 004.8
DOI:

МУЛЯР ЕДУАРД
Хмельницький національний університет
ORCID ID: 0000-0003-4052-4696
e-mail: edikmulyar228@gmail.com
БАГРІЙ РУСЛАН
Хмельницький національний університет
ORCID ID: 0000-0001-5219-1185
e-mail: bahriiro@khmnu.edu.ua
ПАСІЧНИК ОЛЕКСАНДР
Хмельницький національний університет
ORCID ID: 0000-0002-8760-4688
e-mail: o.a.pasichnyk@gmail.com
МАНЗЮК ЕДУАРД
Хмельницький національний університет
ORCID ID: 0000-0002-7310-2126
e-mail: eduard.em.km@gmail.com

МЕТОД ВИЯВЛЕННЯ ЗОВНІШНІХ ПРОЯВІВ НАСИЛЬСТВА У ВІДЕОПОТОЦІ НЕЙРОМЕРЕЖЕВИМИ ЗАСОБАМИ

Проблема виявлення проявів насильства за зображеннями у відеопотоці є актуальною в сучасному світі зі зростаючою кількістю відеоматеріалів, що містять насильницькі сцени. Це включає відео, зняте на вулицях, в громадських місцях та відеозаписи з камер спостереження. Виявлення та реагування на такі сцени важливі для забезпечення безпеки у громадських просторах та захисту прав людини.

Для інтелектуалізації процесу відеоспостереження сьогодні активно використовуються інформаційні технології, а саме нейромережі. Застосування нейромережевих засобів у відеоспостереженні є важливим засобом, оскільки дозволяє автоматично аналізувати великі обсяги відеоматеріалів і виявляти насильницькі сцени з високою точністю.

У статті пропонується метод виявлення зовнішніх проявів насильства за зображеннями у відеопотоці за допомогою згорткової нейронної мережі та класифікатора SVM. На вхід методу подаються кадри відеоматеріалу з яких згорткова нейронна мережа видучає набір ознак, який потім передається класифікатору SVM для отримання оцінки щодо ймовірності належності цих ознак до певного класу (насильницького або не насильницького). Особливістю запропонованого методу є можливість працювати із відеоматеріалом у реальному часі. Це досягається за рахунок того, що згорткова нейронна мережа використовуючи метод fine-tuning навчалася на неперервному потоці даних із мультимедійних платформ для онлайн трансляцій.

Проведено експерименти з використанням різних наборів даних для оцінки ефективності запропонованого методу. Результати показали, що метод досягає високої точності (87,4%-99,45%) виявлення насильства та працює ефективно з відеопотоком даних у реальному часі.

Ключові слова: насильство, виявлення, відеопотік, нейромережі, згорткова нейронна мережа, SVM.

Eduard MULIAR, Ruslan BAHRII, Alexander PASICHNUK, Eduard MANZIUK
Khmelnytskyi National University

METHOD OF DETECTING OUTWARD MANIFESTABONS OF VIOLENCE IN VIDEO STREAMS USING NEURAL NETWORK TOOLS

The problem of detecting violence from images in a video stream is relevant in today's world with a growing number of videos containing violent scenes. This includes video taken on the streets, in public places, and from surveillance cameras. Identifying and responding to such scenes is important for ensuring safety in public spaces and protecting human rights.

Information technologies, namely neural networks, are being actively used to intellectualize the video surveillance process. The use of neural network tools in video surveillance is an important tool, as it allows to automatically analyze large amounts of video materials and detect violent scenes with high accuracy.

The article proposes a method for detecting external manifestations of violence in images in a video stream using a convolutional neural network and an SVM classifier. The input to the method is video frames from which the convolutional neural network extracts a set of features, which is then passed to the SVM classifier to obtain an estimate of the probability of these features belonging to a certain class (violent or non-violent). The peculiarity of the proposed method is the ability to work with video material in real time. This is achieved due to the fact that the convolutional neural network was trained using the fine-tuning method on a continuous stream of data from multimedia platforms for online broadcasts.

Experiments were conducted using different datasets to evaluate the effectiveness of the proposed method. The results showed that the method achieves high accuracy (87,4%-99,45%) in detecting violence and works efficiently with a real-time video data stream.

The use of neural network tools to detect violence in a video stream has great potential in various fields, including public safety, cybersecurity, and human rights protection. Improving the proposed method can help to expand the possibilities of detecting and preventing violence in video streams.

Keywords: violence, detection, video stream, neural networks, convolutional neural network, SVM.

Постановка проблеми

Сьогодні для протидії такій суспільній проблемі як насильство розпочали активно використовувати системи відеоспостережень. Такі країни, як Китай та Південна Корея, є лідерами у встановленні камер відеоспостереження, і результати використання цих систем є вражаючими. У Китаї рівень насильства у громадських місцях знизився на 60%, а в Південній Кореї – на, приблизно, 50% [1].

Однак, ці системи мають деякі недоліки. Основною проблемою є людський фактор, зокрема неуважність та недбалість операторів-спостерігачів. Зазвичай оператор може ефективно контролювати лише обмежену кількість камер відеоспостереження. Однак, коли кількість камер перевищує межі сприйняття, оператори можуть допускати помилки або пропускати випадки насильства.

Для вирішення цієї проблеми сьогодні активно використовуються передові інформаційні технології, зокрема штучні нейронні мережі. Завдяки цим технологіям, системи відеоспостереження можуть автоматично аналізувати великий обсяг відеоданих і виявляти потенційні випадки насильства. Штучний інтелект допомагає зменшити навантаження на операторів і забезпечує більш точне та ефективне виявлення зазначених подій [2].

Застосування штучного інтелекту в системах відеоспостереження є важливим кроком у забезпеченні безпеки у громадських місцях та протидії насильству. Ці технології допомагають забезпечити швидке реагування на зазначені події та вчасне виявлення потенційних загроз. Використання штучного інтелекту в системах відеоспостереження покращує загальний рівень безпеки та сприяє створенню безпечніших громадських просторів.

Аналіз останніх джерел

Для реалізації інтелектуального відеоспостереження сьогодні активно використовують такі моделі нейронних мереж як згорткова та рекурентна [3]. Згорткова нейронна мережа (CNN) є типом нейронних мереж,

які широко використовуються для обробки зображень та роботи з багатомірними даними. Вони були розроблені саме для розпізнавання зображень та виконання завдань комп'ютерного зору [4]. Рекурентна нейронна мережа (RNN) є типом штучних нейронних мереж, які використовуються для моделювання послідовних даних, таких як мовний текст, часові ряди або музика, а також розпізнавання залежностей та шаблонів у цих даних. Основна відмінність RNN від інших типів нейронних мереж полягає в тому, що вона здатна зберігати попередній контекст та використовувати його для обробки наступних вхідних даних [5].

Зазвичай, наведенні моделі нейромереж не використовуються у «чистому» вигляді, як правило вони виступають центральним ядром яке модифікують або доповнюють іншими методами та моделями. Прикладом такого підходу є модель нейронних мереж BiConvLSTM (Bidirectional Convolutional LSTM). Ця модель поєднує в собі два потужних компоненти: двосторонню згорткову мережу (BiConv) та рекурентну нейронну мережу LSTM (Long Short-Term Memory). BiConvLSTM використовується для аналізу послідовних даних, таких як зображення або відео, з метою виявлення шаблонів та залежностей у цих даних. Основна ідея BiConvLSTM полягає в поєднанні двосторонньої згорткової мережі та LSTM для аналізу просторової та часової інформації в послідовних даних. Двостороння згорткова мережа використовується для виділення локальних ознак з різних частин зображення або відео. Вона дозволяє аналізувати дані як вперед (зліва направо) так і назад (справа наліво), тобто нейромережа може бачити як попередні пікселі в зображенні, так і наступні пікселі, що дозволяє мережі краще розуміти, як різні частини зображення пов'язані між собою. Відповідно це дозволяє аналізувати контекст з обох сторін та виявляти шаблони, які можуть бути присутніми в різних частинах послідовних даних [6].

Приклад використання моделі BiConvLSTM наведено у роботі [7], де розглянуто підхід до виявлення насильства у відео за допомогою методу «просторово-часовий кодер». «Просторово-часовий кодер» є методом який побудований за кількома архітектурами (BiConvLSTM, VGG13 [8]) для кодування кожного відеокадру як набору карт функцій. Ці карти функцій потім передаються до BiConvLSTM для подальшого кодування у часовому напрямку відео. Після цього виконується поелементна максимізація кожного з цих кодувань, щоб створити подання всього відео. Це подання передається класифікатору, щоб визначити, чи містить відео насильство. Щодо результатів роботи даного підходу, то для набору даних «Hockey fights» точність склала 96.54%, для набору даних «Violent flows» - 92.18%.

Іншим прикладом використання нейронних мереж для задачі виявлення насильства у відеопотоці є 3D CNN (3D Convolutional Neural Network). Дана модель є згортковою нейронною мережею, що використовується для обробки тримірних даних, таких як відео, медичні зображення або тримірні моделі. Основна ідея 3D CNN полягає у використанні тримірних згорток для виявлення просторових особливостей даних. Вона подібна до звичайних 2D CNN, що використовуються для обробки зображень, але має додатковий третій вимір для роботи з даними [9].

У роботі [10] запропоновано метод до виявлення насильства у відео за допомогою 3D-CNN. 3D-CNN спочатку обробляє кожний кадр відео, використовуючи набори фільтрів для виявлення важливих ознак, таких як рух, форма та колір. Потім 3D-CNN обробляє послідовність кадрів, використовуючи 3D-фільтри. Це дозволяє 3D-CNN виявляти динамічні ознаки насильства, такі як рухи тіла та взаємодії між людьми. Бінарний класифікатор використовується для класифікації насильства або його відсутності. Класифікація здійснюється шляхом застосування логістичної регресії до вихідного тензора 3D-CNN. На рахунок результатів роботи даного методу, то для набору даних «Hockey fights» точність склала 98.3%, для набору даних «Violent flows» - 97.17%.

В наведених методах виявлення насильства у відео не проведено дослідження з відеоматеріалами в реальному часі, що обмежує їх застосування в реальних задачах: автоматичного сповіщення про насильство або покращення роботи операторів спостереження.

Метою роботи є: розробка методу для виявлення зовнішніх проявів насильства у відеопотоці за допомогою нейромережевих засобів у реальному часі. Метод повинен працювати як зі статичним відеоматеріалом (відеоролик) так і з динамічним (відеопотік в реальному часі). Потрібно здійснити аналіз ефективності роботи запропонованого методу на відповідних наборах даних.

Виклад основного матеріалу

Робота запропонованого методу полягає в отриманні ознак насильства з кадрів вхідного відео за допомогою згорткової нейронної мережі і визначення ступеню насильства у відсотковому відношенні у відеопотоці за допомогою SVM (методу опорних векторів). На рис. 1 зображено архітектуру даного методу.

Етап 1 – Отримання кадрів із вхідного відеоматеріалу

Необхідно розбити вхідний відеоматеріал на послідовність кадрів та перетворити кожен кадр у карту зображень.

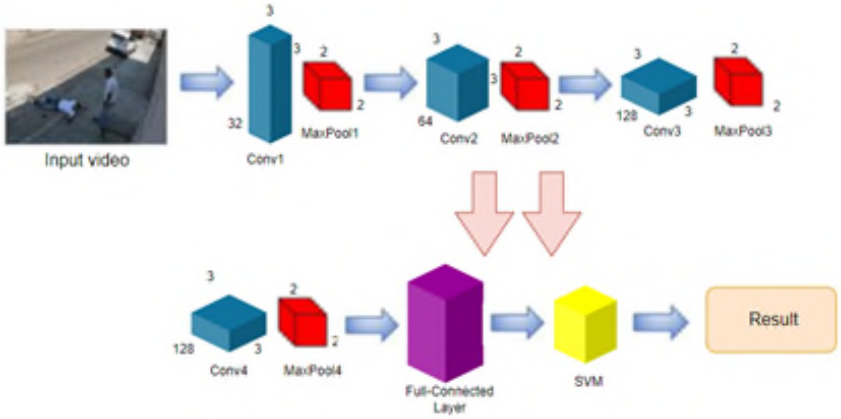


Рис. 1. Артхітектура запропоноваго методу

Етап 2 – Операція згортки

На даному етап необхідно виконати операцію згортки вхідного зображення для отримання карти ознак. Для виконання даної операції використовуються фільтри (матриця параметрів). Для даної нейронної мережі було обрано 4 рівня фільтрів по 32, 64, 128, 128 фільтрів на кожному рівні відповідно. Формування карти ознак можна здійснити за допомогою наступної формули:

$$M(i,j) = (K * X)(i,j) = \sum_{m} \sum_{n} K(m,n) X(i-m, j-n), \quad (1)$$

де M – елемент карти ознак з координатами i та j, X – вхідне зображення, K – детектор ознак, (m, n) – розмірності детектора ознак.

Етап 3 – Операція максимального об'єднання

Максимальне об'єднання є операцією, яка об'єднує елементи в межах фільтра на карті ознак і вибирає найбільший елемент. Тобто, після проходження через шар максимального об'єднання, отримується нова карта ознак, яка містить найбільш помітні ознаки з попередньої карти ознак. Виконати дану операцію можна за допомогою наступної формули:

$$p(i,j) = \max_{i,j} (x(i-m, j-n)), \quad (2)$$

де $p(i, j)$ – значення елементу поточного рівня з координатами i та j, x – вхідні дані з попередніх рівнів, (m, n) – розмірність рецептивного поля.

Етап 4 – Повнозв'язний рівень

Повнозв'язний рівень є моделлю багаторівневого персептрона, де всі нейрони з наступного шару з'єднані з нейронами попереднього шару. Цей рівень використовуються на передостанньому етапі роботи мережі для підготовки результатів на виході мережі. На даному рівні виконується обчислення скалярного добутку даних та параметрів з додаванням зсуву.

Етап 5 – Класифікація отриманих ознак за допомогою SVM

Метод опорних векторів (SVM) використовується для знаходження параметрів гіперплощини у багатомірному просторі, яка може служити для класифікації. Головна ідея полягає в тому, щоб знайти гіперплощину, яка має найбільшу відстань до найближчих навчальних точок будь-якого класу. Ця відстань називається «функціональним запасом». Чим більший функціональний запас, тим менша буде помилка узагальнення класифікатора. На виході класифікатор SVM видає оцінку, яка є ймовірністю того, що вхідні дані належать до певного класу (насильницького або не насильницького).

Алгоритмом, який оновлюватиме вагові коефіцієнти нейромережі під час процесу навчання обрано зворотне поширення похибки. Робота даного алгоритму полягає в наступному: відбувається процес знаходження градієнтів помилок – числових коефіцієнтів (відношення вхідних даних, зсуву до функції втрат), які використовуються для оновлення ваг кожного рівня мережі. Даний алгоритм здійснює оновлення ваг з кінця мережі до початку, у випадку архітектури, що розглядається, від повністю зв'язного рівня до рівня виконання операції згортки.

Для того, щоб запропонований метод міг працювати з відеоматеріалом в реальному часі під час процесу навчання нейронної мережі використовувався метод fine-tuning. Суть даного методу полягає в тому, що нейромережу спочатку навчають на наборі даних готових відео, а потім поступово додають до набору даних відеопотоки у реальному часі. Це допомагає нейромережі навчитися розпізнавати насильство в умовах шуму, швидкої зміни та різноманітності.

Аналіз ефективності запропонованого методу

Для того, щоб оцінити ефективність роботи запропонованого методу, проведено декілька експериментів щодо трьох наборів даних для виявлення насильства: «Hockey fights» [11], «Violent flows» [12], «Livestream» [13]. Оцінка ефективності методу базується на визначенні загальної точності для кожного набору даних.

Для визначення загальної точності роботи методу на відповідному набору даних застосовано наступний підхід:

Етап 1 – Визначення середньої точності

Необхідно зі всіх точностей (accuracy) знайти точність з найбільшим значенням, знаючи цю точність та її епоху потрібно сформувати новий масив даних точностей, які знаходяться в радіусі 10 епох від цієї максимальної точності. Отримавши масив точностей можна знайти середню точність за допомогою наступної формули:

$$A = \frac{\sum_{i=1}^{N} x_i}{N}, \tag{3}$$

де A – середня точність, N – загальна кількість влучень, x – значення відповідної точності, i – порядковий номер.

Етап 2 – Визначення стандартного відхилення

Отримавши середню точність можна знайти стандартне відхилення за наступною формулою:

$$\sigma = \sqrt{\frac{\sum_{i=1}^{N} (x_i - A)^2}{N}}, \tag{4}$$

де σ - стандартне відхилення, N – загальна кількість влучень, x – значення відповідної точності, i – порядковий номер, A – середня точність.

Таким чином, отримане значення середньої точності буде відповідати загальній точності, значення стандартного відхилення буде відповідати похибці середнього значення, яку можна виразити як ± значення, тобто загальна точність = середня точність ± стандартне відхилення.

Загальна точність для набору даних «Hockey fights» склала 98.5%, зображено на рисунку 2. Також у таблиці 1 наведено порівняння точності методів для набору даних «Hockey fights».

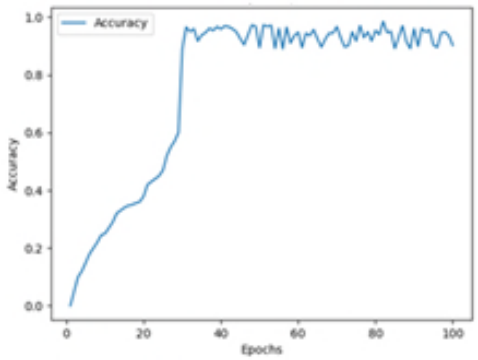


Рис. 2. Загальна точність набору даних «Hockey fights»

Таблиця 1

Порівняння точності методів на наборі даних «Hockey fights»

Метод	Hockey fights
Запропонований	98.5 ± 0.78%
Просторово-часовий кодер [7]	96.54 ± 1.01%
3D CNN [10]	98.3 ± 0.81%

Загальна точність для набору даних «Violent flows» склала 99.45%, зображено на рисунку 3. Наведено у таблиці 2 порівняння точності методів для набору даних «Violent flows».

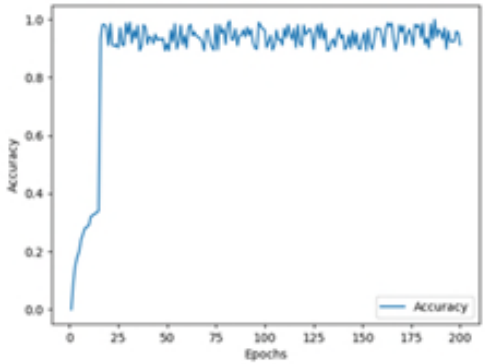


Рис. 3. Загальна точність набору даних «Violent flows»

Таблиця 2

Порівняння точності методів на наборі даних «Violent flows»

Метод	Violent flows
Запропонований	99.45 ± 0.37%

Просторово-часовий кодер [7]	92.18 ± 3.29%
3D CNN [10]	97.17 ± 0.95%

Загальна точність для набору даних «Livestream» склала 87.4%, зображено на рисунку 4. У таблиці 3 зображено порівняння точності методів для набора даних «Livestream».

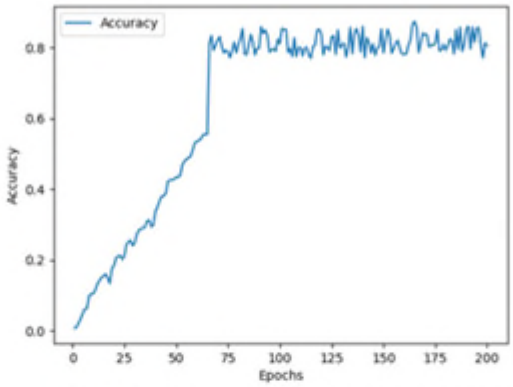


Рис. 4. Загальна точність набору даних «Livestream»

Таблиця 3

Порівняння точності методів на наборі даних «Livestream»

Метод	Livestream
Запропонований	87.4 ± 2.19%
Просторово-часовий кодер [7]	-
3D CNN [10]	-

Особливістю запропонованого методу є робота з відеоматеріалом (відеопотоком) у реальному часі. Дана можливість досягається за рахунок того, що згорткова нейронна мережа навчена на неперервному потоку даних з мультимедійних платформ для онлайн трансляцій використовуючи метод fine-tuning. Тобто, навчання відбувається в режимі реального часу і триватиме, доки примусово не зупиниться трансляція. Відповідно, тестування запропонованого методу відбувалося аналогічним чином, методу на вхід подавалася трансляція і він в реальному часі видавав оцінку сцени, яка відображалася на трансляції.

Висновки

Отже, запропонований метод для виявлення зовнішніх проявів насильства за допомогою нейромережевих засобів дозволяє визначити ступінь насильницького характеру у відсотковому відношенні, на статичних і динамічних відеоматеріалах. Метод на вхід приймає відеоматеріал з якого згорткова нейронна мережа вилучає набір ознак. Потім вилучений набір ознак передається класифікатору SVM, який визначає ймовірність належності вхідних даних до певного класу: насильницького або не насильницького. Головна особливість цього методу полягає в тому, що він може працювати з відеоматеріалом у реальному часі. Це досягається завдяки тому, що згорткова нейронна мережа навчалася на неперервному потоці даних із мультимедійних платформ для онлайн трансляцій за допомогою методу fine-tuning. Проведено експерименти з використанням різних наборів даних для оцінки ефективності запропонованого методу. Результати показали,

що метод досягає високої точності (87,4%-99,45%) виявлення насильства та працює ефективно з відеопотоком даних у реальному часі.

Подальші дослідження спрямовані на пришвидшення роботи та покращення точності запропонованого методу для набору даних, які пов'язані з насильницькими діями у реальному часі.

Література

1. Violence a global public health problem [Електронний ресурс]. – Режим доступу: https://www.scielo.br/j/csc/a/3hrn64cpBqBFb9mNfP4KGXr/?lang=en

2. Використання технологій штучного інтелекту протидії злочинності [Електронний ресурс]. – Режим доступу: https://ivpz.kh.ua/wp-content/uploads/2020/12/Матеріали-семінару_Використання-техн-штучного-інтел_5.11.2020.pdf

3. A CNN-RNN Combined Structure for Real-World Violence Detection in Surveillance Cameras [Електронний ресурс]. – Режим доступу: https://www.mdpi.com/2076-3417/12/3/1021

4. What Is a Convolutional Neural Network? [Електронний ресурс]. – Режим доступу: https://www.ibm.com/topics/convolutional-neural-networks

5. Рекурентна нейронна мережа (RNN): види, навчання, приклади [Електронний ресурс]. – Режим доступу: https://neurohive.io/ru/osnovy-data-science/rekurrentnye-nejronnye-seti/

6. NABNet: A Nested Attention-guided BiConvLSTM network [Електронний ресурс]. – Режим доступу: https://www.sciencedirect.com/science/article/abs/pii/S1746809422007017

7. Convolutional LSTM for the Detection of Violence in Videos [Електронний ресурс]. – Режим доступу: https://openaccess.thecvf.com/content_ECCVW_2018/papers/11130/Hanson_Bidirectional_Convolutional_LSTM_for_the_Detection_of_Violence_in_Videos_ECCVW_2018_paper.pdf

8. Simonyan K., Zisserman A.: Very deep convolutional networks for large-scale image recognition. In International Conference on Learning Representations (2015) [Електронний ресурс]. – Режим доступу: http://arxiv.org/abs/1409.1556

9. 3D Convolutional Neural Network — A Guide for Engineers [Електронний ресурс]. – Режим доступу: https://www.neuralconcept.com/post/3d-convolutional-neural-network-a-guide-for-engineers

10. Jiang X., Xu K., Sun T., Li J.: Efficient Violence Detection Using 3D Convolutional Neural Networks [Електронний ресурс]. – Режим доступу: https://www.researchgate.net/profile/Tanfeng-Sun/publication/337537845_Efficient_Violence_Detection_Using_3D_Convolutional_Neural_Networks/links/5f75e252a6fdcc00864ccb95/Efficient-Violence-Detection-Using-3D-Convolutional-Neural-Networks.pdf

11. Hockey Fight Detection Dataset [Електронний ресурс]. – Режим доступу: https://paperswithcode.com/dataset/hockey-fight-detection-dataset

12. Violent-Flows [Електронний ресурс]. – Режим доступу: https://paperswithcode.com/dataset/violent-flows

13. Livestream [Електронний ресурс]. – Режим доступу: https://www.twitch.tv/directory/category/just-chatting

References

1. Violence a global public health problem [Elektronnyi resurs]. – Rezhym dostupu: https://www.scielo.br/j/csc/a/3hrn64cpBqBFb9mNfP4KGXr/?lang=en

2. Vykorystannia tekhnolohii shtuchnoho intelektu protydii zlochynnosti [Elektronnyi resurs]. – Rezhym dostupu: https://ivpz.kh.ua/wp-content/uploads/2020/12/Матеріали-семінару_Використання-техн-штучного-інтел_5.11.2020.pdf

3. A CNN-RNN Combined Structure for Real-World Violence Detection in Surveillance Cameras [Elektronnyi resurs]. – Rezhym dostupu: https://www.mdpi.com/2076-3417/12/3/1021

4. What Is a Convolutional Neural Network? [Elektronnyi resurs]. – Rezhym dostupu:

https://www.ibm.com/topics/convolutional-neural-networks

5. Rekurentna neironna merezha (RNN): vydy, navchannia, pryklady [Elektronnyi resurs]. – Rezhym dostupu: https://neurohive.io/ru/osnovy-data-science/rekurrentnye-nejronnye-seti/

6. NABNet: A Nested Attention-guided BiConvLSTM network [Elektronnyi resurs]. – Rezhym dostupu: https://www.sciencedirect.com/science/article/abs/pii/S1746809422007017

7. Convolutional LSTM for the Detection of Violence in Videos [Elektronnyi resurs]. – Rezhym dostupu: https://openaccess.thecvf.com/content_ECCVW_2018/papers/11130/Hanson_Bidirectional_Convolutional_LSTM_for_the_Detection_of_Violence_in_Videos_ECCVW_2018_paper.pdf

8. Simonyan, K., Zisserman, A.: Very deep convolutional networks for large-scale image recognition. In International Conference on Learning Representations (2015) [Elektronnyi resurs]. – Rezhym dostupu: http://arxiv.org/abs/1409.1556

9. 3D Convolutional Neural Network — A Guide for Engineers [Elektronnyi resurs]. – Rezhym dostupu: https://www.neuralconcept.com/post/3d-convolutional-neural-network-a-guide-for-engineers

10. Jiang X., Xu K., Sun T., Li J.: Efficient Violence Detection Using 3D Convolutional Neural Networks [Elektronnyi resurs]. – Rezhym dostupu: https://www.researchgate.net/profile/Tanfeng-Sun/publication/337537845_Efficient_Violence_Detection_Using_3D_Convolutional_Neural_Networks/links/5f75e252a6fdcc00864ccb95/Efficient-Violence-Detection-Using-3D-Convolutional-Neural-Networks.pdf

11. Hockey Fight Detection Dataset [Elektronnyi resurs]. – Rezhym dostupu: https://paperswithcode.com/dataset/hockey-fight-detection-dataset

12. Violent-Flows [Elektronnyi resurs]. – Rezhym dostupu: https://paperswithcode.com/dataset/violent-flows

13. Livestream [Elektronnyi resurs]. – Rezhym dostupu: https://www.twitch.tv/directory/category/just-chatting

Додаток В

Презентаційний матеріал

КВАЛІФІКАЦІЙНА РОБОТА МАГІСТРА

Метод виявлення зовнішніх проявів насильства у відеопотоці нейромережевими засобами

Виконав: студент II курсу ОР «Магістр», група КНм-22-1, **Е.Р. Муляр**

Керівник: к.т.н., доцент кафедри КН, **Р.О. Багрій**

Актуальність

Сьогодні для протидії такій суспільній проблемі як насильство розпочали активно використовувати системи відеоспостережень. Такі країни, як Китай та Південна Корея, є лідерами у встановленні камер відеоспостереження, і результати використання цих систем є вражаючими. У Китаї рівень насильства у громадських місцях знизився на 60%, а в Південній Кореї - на приблизно 50%.

Однак, ці системи мають деякі недоліки. Основною проблемою є людський фактор, зокрема неуважність та недбалість операторів-спостерігачів. Зазвичай оператор може ефективно контролювати лише обмежену кількість камер відеоспостереження. Однак, коли кількість камер перевищує їх межі, оператори можуть допускати помилки або пропускати випадки насильства.

Для вирішення цієї проблеми сьогодні активно використовуються передові інформаційні технології, зокрема штучні нейронні мережі. Завдяки цим технологіям, системи відеоспостереження можуть автоматично аналізувати великий обсяг відеоданих і виявляти потенційні випадки насильства. Штучний інтелект допомагає зменшити навантаження на операторів і забезпечує більш точне та ефективне виявлення подій.

Мета роботи

Метою кваліфікаційної роботи магістра є розробка методу виявлення зовнішніх проявів насильства у відеопотоці нейромережевими засобами.

Зокрема, запропонований метод повинен виявляти прояви насильства на неперервному відеопотоці даних у реальному часі (прямі відеотрансляції).

Завдання

Задля досягнення поставленої мети визначені наступні завдання роботи:

- Провести аналіз нейромережевих моделей та існуючих підходів для виявлення проявів насильства у відеопотоці;
- Розробити метод виявлення зовнішніх проявів насильства у відеопотоці з використанням згорткової нейронної мережі та класифікатора SVM;
- Підготувати набір даних для навчання згорткової нейронної мережі;
- Навчити попередньо навчену згорткову нейронну мережу виявляти ознаки насильства на неперервному відеопотоці даних;
- Визначити загальну точність запропонованого методу виявлення зовнішніх проявів насильства.

Об’єкт та предмет дослідження

Об’єкт дослідження – процес виявлення прояву насильства у відеопотоці нейромережевими засобами.

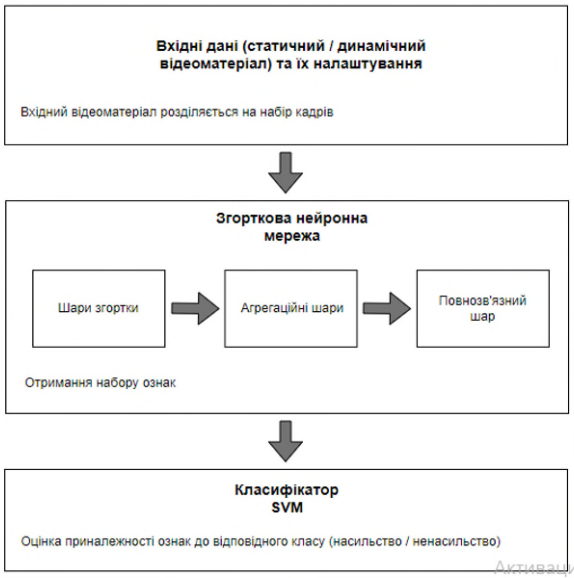
Предмет дослідження – моделі нейронної мережі, методи класифікації ознак для виявлення прояву насильства у відеопотоці.

Наукова новизна

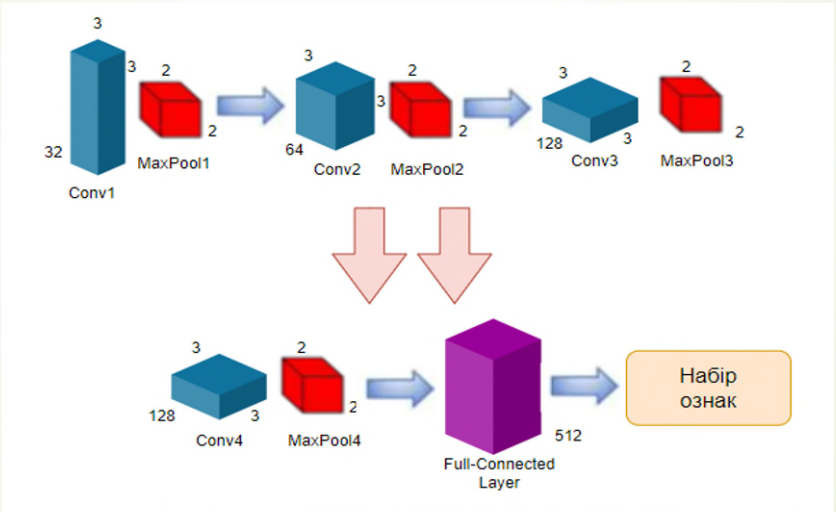
В результаті проведеної роботи були отримані наступні результати:

- вдосконалено архітектуру згорткової нейронної мережі, що дало можливість виявляти прояви насильства у відеопотоці даних у реальному часі, що досягається за рахунок додаткового навчання попередньо навченої моделі на неперервному відеопотоці даних який містить прояви насильства;
- розроблено метод виявлення зовнішніх проявів насильства у відеопотоці за допомогою згорткової нейронної мережі та класифікатора SVM, що дозволило підвищити точність виявлення проявів насильства до 87.4%-99.45% у неперервному відеопотоці у реальному часі.

Схема методу виявлення зовнішніх проявів насильства у відеопотоці нейромережевими засобами



Архітектура згорткової нейронної мережі в контексті запропонованого методу

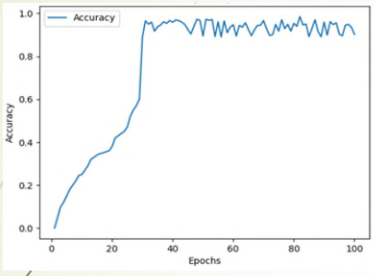


Навчання згорткової нейронної мережі

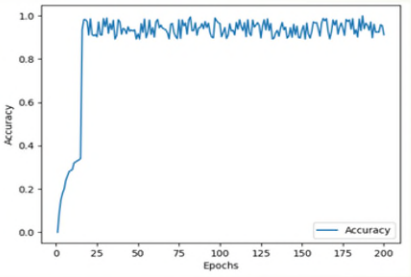
Зворотне поширення представляє собою ітеративний метод навчання суть якого полягає у знаходженні градієнту кожного нейрона по відношенню до функції втрат, щоб визначити, наскільки вихідні дані вносять вклад в загальні втрати. В даному методі процес навчання відбувається з кінця, тобто поширення сигналів помилки йдуть від виходів мережі до її входів, з точки зору представленої архітектури від повністю зв'язаного шару до шару згортки.

Fine-tuning є процесом використання попередньо навченої моделі машинного навчання і подальшого навчання цієї моделі на нових даних для покращення її точності на конкретній задачі. Основна ідея застосування fine-tuning полягає в тому, що попередньо навчена модель вже має досвід вирішення загальних завдань, у випадку поставленої задачі це виявлення ознак насильства на статичних відеоматеріалах. Цей досвід може бути використаний, щоб покращити точність моделі на задачі виявлення зовнішніх проявів насильства у прямій відеотрансляції (динамічний відеоматеріал).

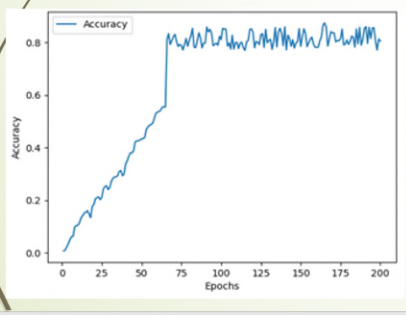
Аналіз ефективності роботи методу



Точність тесту для набору даних « Hockey Fights» має загальну точність 98.5 ± 0.78%

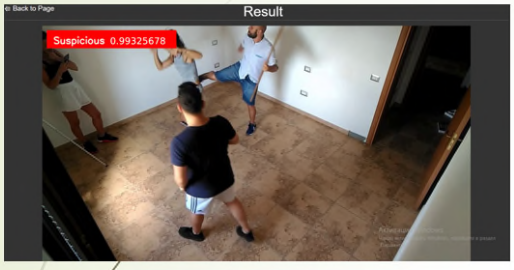


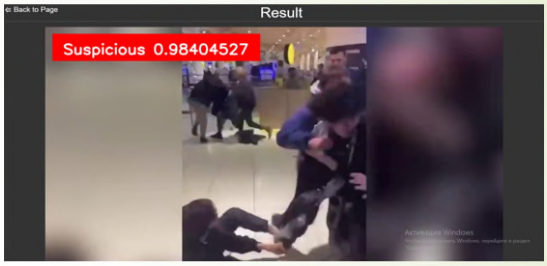
Точність тесту для набору даних «Violent Flows» становить 99.45 ± 0.37%.

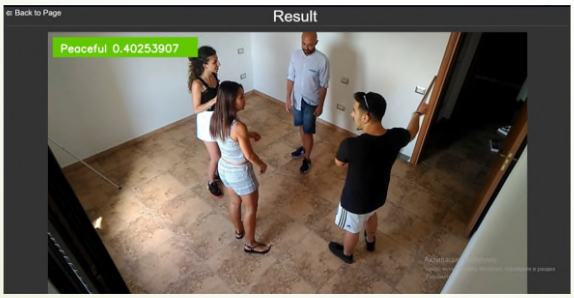


Точність тесту для набору даних «Livestream» становить 87.4 ± 2.19%.

Приклад роботи запропонованого методу







Висновки

В результаті кваліфікаційної роботи магістра розроблено метод виявлення зовнішніх проявів насильства у відеопотоці нейромережевими засобами.

В процесі розробки методу виконано наступні задачі:

- Досліджено предметну область поставленого завдання та проаналізовано існуючі підходи та публікації з метою доведення актуальності проблеми, яку необхідно вирішити.
- Розроблено метод виявлення зовнішніх проявів насильства у відеопотоці нейромережевими засобами. Запропонований метод складається з 2 частин: згорткова нейронна мережа (задача полягає у вилученні та формуванні набору ознак) та SVM (задача полягає у розрахуванні оцінки щодо ймовірності належності вхідних даних (ознак) до певного класу (насильницького або не насильницького)). Для того, щоб метод міг працювати з динамічним відеоматеріалом (пряма відеотрансляція) використано метод тонкого налаштування (fine-tuning).
- Проведено валідацію розробленого методу на тестових наборах даних: «Hockey fights» та «Violent flows» (в якості статичного відеоматеріалу), «Livestream» (в якості динамічного відеоматеріалу). Для набору даних «Hockey fights» загальна точність становить 98.5±0.78%, для «Violent flows» склала 99.45±0.37%, для «Livestream» відповідно 87.4±2.19%.


Дякую за
увагу!
Активация Windows

Додаток Г

Програмні коди

```
//model_train.ipynb
# Import Stuff
import tensorflow as tf
from tensorflow import keras
import numpy as np
from tensorflow.keras import datasets, layers, models
from sklearn.model_selection import train_test_split
import matplotlib.pyplot as plt
import cv2
from imgextract import Extractor
from tensorflow.keras.preprocessing.image import
ImageDataGenerator
from multiprocessing import Process
from IPython.display import clear_output
# All Parameters required for training are declared
over here
# Frequency of Image Capturing
FRAME_SKIP = 2
# Frame Size
FRAME_SIZE = (150,150)
# Dataset
!rm -r A-Dataset-for-Automatic-Violence-Detection-in-
Videos/
!git clone https://github.com/airtlab/A-Dataset-for-
Automatic-Violence-Detection-in-Videos
!rm -r Data
!mkdir Data
!mkdir -p ./Data/Video/Violent
!mkdir -p ./Data/Video/NonViolent
!cp -a ./A-Dataset-for-Automatic-Violence-Detection-
in-Videos/violence-detection-dataset/violent/cam1/.
./Data/Video/Violent/
!cp -a ./A-Dataset-for-Automatic-Violence-Detection-
in-Videos/violence-detection-dataset/non-
violent/cam1/. ./Data/Video/NonViolent/
clear_output()
!mkdir -p ./Data/Training/V
!mkdir -p ./Data/Training/NV
def thread_1():
    ext = Extractor(FRAME_SIZE, FRAME_SKIP)
    for i in range(60):
        path = f"./Data/Video/Violent/{i+1}.mp4"
        print(f"Processing Violent Vid-{i}")
        ext.extract(path, 'V')
    print("Violent Extracted")
def thread_2():
    ext = Extractor(FRAME_SIZE, FRAME_SKIP)
    for i in range(60):
        path =
f"/content/Data/Video/NonViolent/{i+1}.mp4"
        print(f"Processing NonViolent Vid-{i}")
        ext.extract(path, 'NV')
    print("Non-Violent Extracted")
# Violent Extraction
t1 = Process(target=thread_1, args=())
t2 = Process(target=thread_2, args=())

t1.start()
t2.start()
# NonViolent Extraction

t1.join()
t2.join()
print("Complete")
base_dir='./Data'
train_dir=os.path.join(base_dir,'Training')
train_violent_dir =os.path.join(train_dir, 'V' )
train_nonviolent_dir=os.path.join(train_dir,'NV')
train_datagen= ImageDataGenerator(rescale=1./255,
rotation_range=40,width_shift_range=0.2,
height_shift_range=0.2,
shear_range=0.2,horizontal_flip=True,
fill_mode='nearest')
train_generator =
train_datagen.flow_from_directory(train_dir,color_mod
e="rgb", target_size =
FRAME_SIZE,batch_size=20,classes=['NV','V'],
class_mode='binary', shuffle=True)
model= tf.keras.models.Sequential([
      tf.keras.layers.Conv2D(32,(3,3),activation='re
lu',input_shape=(150,150,3)),
      tf.keras.layers.MaxPooling2D(2,2),
      tf.keras.layers.Conv2D(64,(3,3),activation='re
lu'),
      tf.keras.layers.MaxPooling2D(2,2),
      tf.keras.layers.Conv2D(128,(3,3),activation='r
elu'),
      tf.keras.layers.MaxPooling2D(2,2),
```

```
        tf.keras.layers.Conv2D(128,(3,3),activation='relu'),
        tf.keras.layers.MaxPooling2D(2,2),
        tf.keras.layers.Dropout(0.5),
        tf.keras.layers.Flatten(),
        tf.keras.layers.Dense(512, activation='relu'),
        tf.keras.layers.Dense(1,activation ='sigmoid')

])

# Freeze the layers of the pre-trained model
for layer in base_model.layers:
    layer.trainable = False
model.compile(loss='binary_crossentropy',optimizer='adam',metrics=['accuracy'])
model1=model.fit(train_generator,steps_per_epoch=50, epochs=30)
# SVM layer
svm_model = models.Sequential([
    model,
    layers.Dense(1, activation='linear')  # Linear activation for SVM
])

# Compile the SVM model
svm_model.compile(optimizer='adam',
                  loss='hinge',  # Hinge loss for SVM
                  metrics=['accuracy'])

# Train the SVM model
svm_model.fit(train_generator, steps_per_epoch=50, epochs=10)  # Adjust the number of epochs as needed

# Save the final model


import time
t = time.time()
export_path_keras = "./{}.h5".format(int(t))
print(export_path_keras)


svm_model.save(export_path_keras)

//views.py
from django.views.decorators.clickjacking import xframe_options_exempt
import numpy as np
from django.shortcuts import render, redirect, reverse
from django.http import JsonResponse, HttpResponse, StreamingHttpResponse
from django.views.decorators.csrf import csrf_exempt
from .models import DocModel
import json
from django.views.decorators import gzip
import cv2
from .forms import DocumentForm
from django.conf import settings
model = settings.MODEL


class VideoCamera(object):
    def __init__(self, url=None):
        self.font = cv2.FONT_HERSHEY_SIMPLEX
        self.status = True
        self.org = (50, 80)
        self.fontScale = 1.4
        self.thickness = 3
        self.SIZE = (150, 150)
        self.THRESH = 0.5
        self.url = 0 if url is None else '.'+url
        self.video = cv2.VideoCapture(self.url)
        self.skipCount = 2
        self.prev = None
        self.fcount = 0

    def __del__(self):
        self.video.release()

    def get_frame(self):
        ret, image = self.video.read()
        if not ret:
            self.status = False
            pass

        if self.fcount % self.skipCount == 0:
            tmp = cv2.resize(image, self.SIZE)
            tmp = tmp / 255.0
            pred = model.predict(np.array([tmp]))
            string = "Suspicious" if pred[0][0] > self.THRESH else "Peaceful"
            string += f" {str(pred[0][0])}"
            self.prev = string

        else:
            string = self.prev

        color = (255, 255, 255)
```

```
        image = cv2.rectangle(image, (20, 20), (600,
100), (0, 200, 100), cv2.FILLED) if string.split('
')[0] == 'Peaceful' else cv2.rectangle(
            image, (20, 20), (600, 100), (0, 0, 255),
cv2.FILLED)
        image = cv2.putText(image, string, self.org,
self.font,
                            self.fontScale, color,
self.thickness, cv2.LINE_AA)
        ret, jpeg = cv2.imencode('.jpg', image)
        self.fcount += 1
        return jpeg.tobytes()


def gen(camera):
    while camera.status:
        frame = camera.get_frame()
        yield(b'--frame\r\n'b'Content-Type:
image/jpeg\r\n\r\n' + frame + b'\r\n\r\n')


@gzip.gzip_page
def Stream(request):
    try:
        entry = DocModel.objects.all().last()
        return
StreamingHttpResponse(gen(VideoCamera(entry.vid.url))
, content_type="multipart/x-mixed-
replace;boundary=frame")
    except
StreamingHttpResponse.HttpResponseServerError as e:
        print("aborted")


@gzip.gzip_page
def StreamToken(request, token):
    try:
        entry =
DocModel.objects.filter(stoken=token).last()
        return
StreamingHttpResponse(gen(VideoCamera(entry.vid.url))
, content_type="multipart/x-mixed-
replace;boundary=frame")
    except
StreamingHttpResponse.HttpResponseServerError as e:
        print("aborted")


def HomeView(request):
    if request.method == 'POST':
        form = DocumentForm(request.POST,
request.FILES)
        if form.is_valid():
            form.save()
            return redirect('streamroom')

    else:
        form = DocumentForm()
        return render(request, 'home.html', {'form':
form})


# @xframe_options_exempt
def StreamView(request):
    entry = DocModel.objects.all().last()
    if entry is None:
        return JsonResponse({'message': 'No Video
Files Yet!'})
    return render(request, 'stream.html')


# API End Point
def StreamTokenView(request, token):
    try:
        entry =
DocModel.objects.filter(stoken=token).last()
        if entry is None:
            return JsonResponse({'message': 'Token
Not Registered'})

        return render(request, 'streamtoken.html',
{'token': token})

    except DocModel.DoesNotExist:
        return JsonResponse({'message': 'Token Not
Registered'})


@csrf_exempt
def APIEnd(request):
    if request.method == 'POST':
        try:
            stoken = request.POST['stoken']
            vidFile = request.FILES['vid']
            DocModel(stoken=stoken,
vid=vidFile).save()
            baseurl =
request.build_absolute_uri(reverse('home'))
```

```
            return JsonResponse({'status': 'ok',
'message': f'Files Received from sender {stoken}',
'vidurl': baseurl+'streamtoken/'+stoken})
        except:
            return HttpResponse(status=400)

    return JsonResponse({'status': 'Wait kro bhai'})
# Create your views here
```

```
//models.py
from django.db import models
import os


def content_file_name(instance, filename):
    ext = filename.split('.')[-1]
    filename = "%s_%s.%s" % (instance.user.id,
instance.questid.id, ext)
    return os.path.join('uploads', filename)


class DocModel(models.Model):
    stoken = models.CharField(max_length=50,
unique=False, default='')
    date = models.DateTimeField(auto_now_add=True)
    vid = models.FileField(upload_to='documents/')
    def __str__(self):
        return str(self.date)
```

```
//urls.py
from django.urls import path
from .views import *
from django.conf import settings
from django.conf.urls.static import static

urlpatterns = [
    path('', HomeView, name='home'),
    path('getstream', Stream, name='streamdt'),
    path('gettokenstream/<token>', StreamToken,
name='streamtk'),
    path('stream/', StreamView, name='streamroom'),
    path('streamtoken/<token>', StreamTokenView,
name='stokenview'),
    path('api/', APIEnd, name='api')
]

if settings.DEBUG:
    urlpatterns += static(settings.MEDIA_URL,

document_root=settings.MEDIA_ROOT)
```

Mon Dec 04 16:10:04 EET 2023, Петровський Сергій Степанович, Хмельницький національний університет, ХНУ

Anti-Plagiarism v-15.257

Максимальне співпадіння з одним документом 1.0%

Словники перевірки: en_US, ru_RU, ua_UA. **Помилок в документах: 11%**

	Документ		Сумарний збіг по Базі Даних	
ID: 121725 Назва: КВАЛІФІКАЦІЙНА РОБОТА МАГІСТРА на тему Метод виявлення зовнішніх проявів насильства у відеопотоці нейромережевими засобами Додано в БД: 2023-12-04 Автора: Е.Р. Муляр Керівники: Р.О. Багрій Консультанти: Опоненти:	Символи	Лексеми	Символи	Лексеми
	86461	1299	2583 (3%)	41 (3%)

Джерело плагіату

ID	Опис	Наявність плагіату в документі	
		Символи	Лексеми

UNICHECK
by Turnitin

Ім'я користувача:
Кафедра КН

ID перевірки:
1015968439

Дата перевірки:
04.12.2023 17:07:47 EET

Тип перевірки:
Doc vs Internet + Library

Дата звіту:
04.12.2023 17:12:26 EET

ID користувача:
100005671

Назва документа: **КНм-22-1 Муляр**

Кількість сторінок: **79** Кількість слів: **14156** Кількість символів: **107314** Розмір файлу: **2.35 MB** ID файлу: **1015647281**

Виявлено модифікації тексту (можуть впливати на відсоток схожості)

10.1%
Схожість

Найбільша схожість: **1.74%** з джерелом з Бібліотеки (ID файлу: **1013023335**)

9.59% Джерела з Інтернету	793	Сторінка 81
3.64% Джерела з Бібліотеки	97	Сторінка 88

0% Цитат

Вилучення цитат вимкнене

Вилучення списку бібліографічних посилань вимкнене

0%
Вилучень

Немає вилучених джерел

Модифікації

Виявлено модифікації тексту. Детальна інформація доступна в онлайн-звіті.

Підозріле форматування	17 сторінок

РІШЕННЯ ЕКСПЕРНОЇ КОМІСІЇ

КАФЕДРИ КОМП'ЮТЕРНИХ НАУК

ПРО ДОПУСК КВАЛІФІКАЦІЙНОЇ РОБОТИ МАГІСТРА ДО ЗАХИСТУ ЗА РЕЗУЛЬТАТАМИ АНАЛІЗУ ЗВІТУ ПОДІБНОСТІ

Підтверджуємо ознайомлення з результатом звіту подібності щодо роботи, генерованого системою виявлення текстових збігів/ідентичності/схожості:

Назва: *Метод виявлення зовнішніх проявів насильства у відеопотоці нейромережевими засобами*

Автор: *Муляр Едуард Русланович*

Спеціальність: 122 – Комп'ютерні науки

Освітня програма: освітньо-професійна

Науковий керівник: к.т.н., доц. Багрій Р.О.

Після аналізу звіту подібності зроблено такий висновок:

№	Висновок	Позначка про відповідність
1	Запозичення, виявлені в роботі, є законними і не є плагіатом. Робота приймається до захисту.	відповідає
2	Виявлені запозичення не є плагіатом, розміщені в розділах, які не описують безпосередньо авторське дослідження, але кількість цитат перевищує обсяг, виправданий поставленою метою роботи Робота приймається до захисту, але має бути відкоригована. Відкоригований варіант має бути поданий на кафедру за 2 дні до захисту, разом із заявою щодо самостійності виконання письмової роботи та ідентичності друкованої та електронної версії роботи	
3	Виявлені запозичення не є плагіатом, але частково розміщені в розділах, які описують безпосередньо авторське дослідження, а кількість цитат перевищує обсяг, виправданий поставленою метою роботи. В зв'язку з цим мета роботи та поставлені завдання не були досягнені. Робота може бути допущена до захисту (наступного року) після того як буде відкоригована та допрацьована і успішно пройде повторну перевірку на академічний плагіат.	
4	Робота містить навмисні текстові спотворення, передбачувані спроби укриття запозичень або інші прояви академічного плагіату. Робота містить фабрикацію або фальсифікацію даних. Робота не допускається до захисту.	

Підтвердження:

Запозичення, виявлені в роботі, є законними і не є плагіатом, оскільки:

1) за програмою Anti-Plagiarism виявлені 1% є фрагментарними – містять поширені конструкції, загальновідомі терміни, скорочення та визначення.
2) За програмою UNICHECK виявлені 10.1%, що є запозиченнями, які розміщені в розділах аналізу існуючих технологій та прототипів, які не описують безпосередньо авторське дослідження і не стосуються результатів роботи.

Сумарний обсяг всіх запозичень, визначений системою виявлення збігів/ідентичності/схожості, складає 1% і 10.1% відповідно, що, з урахуванням наведених обґрунтувань, відповідає характеру наукового дослідження і свідчить на користь кваліфікаційної роботи.

Керівник роботи _____ Руслан БАГРІЙ

Гарант ОП _____ Руслан БАГРІЙ

Завідувач кафедри КН _____ Олександр БАРМАК

ВІДГУК ОПОНЕНТА

на кваліфікаційну роботу магістра

гр. КНм-22-1 Муляра Едуарда Руслановича за темою: *Метод виявлення зовнішніх проявів насильства у відеопотоці нейромережевими засобами*

1. Актуальність обраної теми

Використання штучних нейронних мереж з метою виявлення зовнішніх проявів насильства у відеопотоці дозволить системам відеоспостереження автоматично аналізувати великий обсяг відеоданих і виявляти потенційні випадки насильства, пришвидшити процес виявлення ознак насильства на великій площі спостереження, а також забезпечить усунення негативного людського фактору під час відеоспостереження – неуважність. Тому робота, виконана автором є актуальною та перспективною.

2. Відповідність роботи предметній області спеціальності 122 Комп'ютерні науки та загальним вимогам до наукових робіт

Обрана тема задачі виявлення зовнішніх проявів насильства у відеопотоці нейромережевими засобами, в межах якої виконані поставлені задачі, повною мірою відповідає предметній області спеціальності 122 «Комп'ютерні науки» та вимогам до кваліфікаційної роботи магістра.

3. Повнота розкриття мети та завдань дослідження

В роботі автор повністю розкриває мету дослідження та поставленні в межах теми завдання.

4. Наявність наукової новизни

В кваліфікаційній роботі представлена наукова новизна та інновації, відповідна спеціальності 122 «Комп'ютерні науки» в межах обраної області дослідження. Продемонстровано й обгрунтовано результати, які мають наукове та інноваційне значення. Результати дослідження оприлюднені на науково-практичній конференції «АПКН-2023» та у віснику Хмельницького національного університету

5. Зміст кожного розділу роботи

Робота містить чотири розділи. У першому розділі виконано аналіз сучасного стану проблеми виявлення зовнішніх проявів насильства у відеопотоці нейромережевими засобами. Другий розділ присвячено розробці математичної моделі методу виявлення зовнішніх проявів насильства у відеопотоці нейромережевими засобами. У третьому розділі

виконано розробку веб-додатку методу виявлення зовнішніх проявів насильства у відеопотоці нейромережевими засобами. У четвертому розділі виконано дослідження ефективності методу виявлення зовнішніх проявів насильства у відеопотоці нейромережевими засобами.

6. Ступінь розкриття теми роботи

Тема кваліфікаційної роботи повною мірою розкрита та обгрунтована, проведено аналіз актуальності та відомих досліджень в межах обраної теми, поставлені завдання, які у роботі виконані, та проведено аналіз результатів прикладного застосування запропонованих методу і засобів.

7. Якість оформлення кваліфікаційної роботи

Оформлення роботи відповідає необхідним нормам та вимогам, які ставляться до оформлення кваліфікаційних робіт.

8. Недоліки кваліфікаційної роботи

Рекомендовано розширити перелік скорочень та відкоригувати список використаних джерел згідно сучасних вимог.

Доцільно було б навести приклади тестування запропонованого методу на відеоматеріалах з поганою якістю та не достатньою кількість освітлення.

9. Загальний висновок (допускається чи не допускається до захисту), якої оцінки заслуговує кваліфікаційна робота

Враховуючи високий рівень виконання та забезпечення усіх необхідних вимог, робота може бути допущена до захисту. Рекомендована оцінка *відмінно*.

Опонент Зав. каф. АКІТтаР Валерій Мартинюк

ХМЕЛЬНИЦЬКИЙ НАЦІОНАЛЬНИЙ УНІВЕРСИТЕТ
МОН УКРАЇНИ
кафедра комп'ютерних наук

ВІДГУК НАУКОВОГО КЕРІВНИКА

на кваліфікаційну роботу магістра

гр. КНм-22-1 Муляра Едуарда Руслановича за темою: Метод виявлення зовнішніх проявів насильства у відеопотоці нейромережевими засобами

1. Актуальність теми

Актуальність теми достатньо обгрунтована, оскільки виявлення зовнішніх проявів насильства у відеопотоці нейромережевими засобами може сприяти зменшенню насильства в суспільстві та забезпечити більш ефективну боротьбу з цією проблемою. Однією з переваг цього методу є можливість автоматичного виявлення насильства на великій кількості відео, що дозволяє економити час та зусилля людей, які займаються цією проблемою.

2. Відповідність роботи предметній області спеціальності 122 Комп'ютерні науки та загальним вимогам до наукових робіт

Теми кваліфікаційної роботи "Метод виявлення зовнішніх проявів насильства у відеопотоці нейромережевими засобами" відповідає предметній області спеціальності 122 Комп'ютерні науки та вимогам до кваліфікаційної роботи магістра, оскільки об'єктом дослідження є процес виявлення прояву насильства у відеопотоці нейромережевими засобами, а саме використання алгоритмів машинного навчання для виявлення певних ознак насильства, предметом дослідження – моделі нейронної мережі, методи класифікації ознак для виявлення прояву насильства у відеопотоці.

3. Професійні та особистісні якості магістранта

Муляр Е. Р. під час роботи над кваліфікаційною роботою магістра продемонстрував високий рівень знань та умінь за спеціальністю "Комп'ютерні науки".

4. Ступінь самостійності під час виконання кваліфікаційної роботи

Робота виконана самостійно, академічного плагіату не виявлено, стосовно всіх запозичень наведено відповідні посилання на джерела.

5. Наукова новизна та оригінальність запропонованих підходів

Отримані такі результати: вдосконалено архітектуру згорткової нейронної мережі, що дало можливість виявляти прояви насильства у відеопотоці даних у реальному часі, що досягається за рахунок додаткового навчання попередньо навченої моделі на неперервному відеопотоці даних який містить прояви насильства; розроблено метод виявлення зовнішніх проявів насильства у відеопотоці за допомогою згорткової нейронної мережі та

класифікатора SVM, що дозволило підвищити точність виявлення проявів насильства до 87.4%-99.45% у неперервному відеопотоці у реальному часі. Отримані результати оприлюднені на XIII всеукраїнської науково-практичної конференції «Актуальні проблеми комп'ютерних наук АПКН-2023», 17-18 листопада 2023 р., м. Хмельницький, Україна, доповідь на тему «Метод виявлення ознак насильства у відеоматеріалах нейромережевими засобами».

6. Ступінь оволодіння методами дослідження

Студент Муляр Е. Р. має достатньо високий ступінь володіння методами дослідження, що були використанні у роботі.

7. Повнота та якість розкриття теми роботи

Мета роботи повністю розкрита, отримані результати підтверджують достовірність наукових положень.

8. Логічність, послідовність, аргументованість, літературна грамотність викладу матеріалу

Викладення матеріалу логічне, послідовне та аргументоване. Мова і стиль викладення кваліфікаційної роботи магістра відповідають стандартам, що забезпечує доступність сприймання матеріалу і відповідає вимогам до сучасних наукових робіт.

9. Можливість практичного застосування кваліфікаційної роботи, окремих її частин

Може мати практичне значення при виявленні проявів насильства при відеоспостереженні в громадських місцях та запобіганні конфліктам, що покращить загальну безпеку.

10. Висновок про можливість допуску кваліфікаційної роботи до захисту, на яку оцінку заслуговує робота

Вважаю, що кваліфікаційна робота студента Муляра Едуарда Руслановича може бути рекомендована до захисту та заслуговує на оцінку "відмінно".

Науковий керівник _____ к.т.н., доц. Руслан Багрій