ADULT CONTENT FILTERING USING MACHINE LEARNING

MSc Research Project
MSc in Cybersecurity

Mahavir Gala
Student ID: 22208208

School of Computing
National College of Ireland

Supervisor: Joel Aleburu

National College of Ireland

Project Submission Sheet

Student Name:	Mahavir Gala
Student ID:	22208208
Programme:	Master in Cybersecurity	**Year:**	2023 - 2024
Module:	MSc Research Project
Lecturer:	Joel Aleburu
Submission Due Date:	11/08/2024
Project Title:	CA-2
Word Count:	……………………………………………………………………………………………… ……………………

I hereby certify that the information contained in this (my submission) is information pertaining to research I conducted for this project. All information other than my own contribution will be fully referenced and listed in the relevant bibliography section at the rear of the project.
ALL internet material must be referenced in the references section. Students are encouraged to use the Harvard Referencing Standard supplied by the Library. To use other author's written or electronic work is illegal (plagiarism) and may result in disciplinary action. Students may be required to undergo a viva (oral examination) if there is suspicion about the validity of their submitted work.

Signature: Mahavir Gala

Date: 12/08/2024

PLEASE READ THE FOLLOWING INSTRUCTIONS:

1. Please attach a completed copy of this sheet to each project (including multiple copies).
2. Projects should be submitted to your Programme Coordinator.
3. **You must ensure that you retain a HARD COPY of ALL projects**, both for your own reference and in case a project is lost or mislaid. It is not sufficient to keep a copy on computer. Please do not bind projects or place in covers unless specifically requested.
4. You must ensure that all projects are submitted to your Programme Coordinator on or before the required submission date. Late submissions will incur penalties.
5. All projects must be submitted and passed in order to successfully complete the year. Any project/assignment not submitted will be marked as a fail.

Office Use Only	
Signature:	
Date:	
Penalty Applied (if applicable):	

AI Acknowledgement Supplement

[Insert Module Name]

[Insert Title of your assignment]

Your Name/Student Number	Course	Date

This section is a supplement to the main assignment, to be used if AI was used in any capacity in the creation of your assignment; if you have queries about how to do this, please contact your lecturer. For an example of how to fill these sections out, please click here.

AI Acknowledgment

This section acknowledges the AI tools that were utilized in the process of completing this assignment.

Tool Name	Brief Description	Link to tool

Description of AI Usage

This section provides a more detailed description of how the AI tools were used in the assignment. It includes information about the prompts given to the AI tool, the responses received, and how these responses were utilized or modified in the assignment. **One table should be used for each tool used**.

[Insert Tool Name]	
[Insert Description of use]	
[Insert Sample prompt]	[Insert Sample response]

Evidence of AI Usage

This section includes evidence of significant prompts and responses used or generated through the AI tool. It should provide a clear understanding of the extent to which the AI tool was used in the assignment. Evidence may be attached via screenshots or text.

Additional Evidence:

[Place evidence here]

Additional Evidence:

[Place evidence here]

Abstract

This research addresses the problem of filtering adult content online, concentrating on the limits of traditional methods as well as looking into sophisticated artificial intelligence techniques. Traditional content filtering methods, including domain-based and keyword-based blocking, are actually usually poor because of their static attributes and failure to adjust to brand-new content. This research study looks into the application of Convolutional Neural Networks (CNNs), particularly the VGG16 architecture, for real-time, correct classification of adult content. CNNs give considerable renovations in detecting subtle patterns as well as conforming to evolving content. Furthermore, the study highlights the importance of integrating vulnerability scanning to attend to security risks related to adult content websites. Through CNN-based filtering along with real-time vulnerability scanning, this work targets to improve online safety and deliver a comprehensive solution for shielding users from inappropriate content and security threats. The findings recommend promising instructions for future analysis as well as technological advancements in content moderation and cybersecurity.

Acknowledgement

I would like to express my deep sense of gratitude to my respected professor for guiding me and helping me to develop this research work. I am very thankful to the university for giving me the chance to work on this topic and gain invaluable knowledge. Lastly, I would like to thank my friends and family for motivating me through this journey.

List of Keywords

- Adult Content Detection
- Convolutional Neural Networks (CNNs)
- VGG16
- Image Classification
- Machine Learning
- Data Augmentation
- Transfer Learning
- Content Filtering
- TensorFlow
- Keras
- Google Colab
- Image Preprocessing
- Model Evaluation
- Feature Extraction
- Content Moderation

Table of Contents

List of Figures

Chapter 1: Introduction

1.1 Research Background

The expansion of the internet has actually boosted accessibility to varied material, including adult material, giving considerable security and protection risks. Traditional filtering approaches, centring only on blocking out explicit content, fall short in dealing with the affiliated safety and security risks, such as malware and phishing (Choudrie *et al.* 2021). This research highlights the demand for an extensive option combining equipment learning-based web content filtering and vulnerability scanning. By leveraging convolutional semantic networks (CNNs) like VGG16 for exact information distinction and incorporating real-time vulnerability detection, the research study targets boosting online protection. This double strategy ensures customers are safeguarded from improper content and surveillance dangers, providing a sturdy safeguard for vulnerable groups (Jani *et al.,* 2020).

The VGG16 architecture, a deep-discovery model renowned for its performance in image recognition, is well-suited for this purpose. Through leveraging VGG16 for categorizing graphics as grown-up or even non-adult content, it is achievable to develop a device that can dynamically and properly filter improper material (Boongasame *et al.* 2023). This method addresses the limitations of traditional methods by concentrating on the content itself as opposed to counting on fixed listings or even domain name limitations (Paullada *et al.,* 2021). Integrating such a CNN-based model into a web extension allows for real-time monitoring as well as filtering, making certain that users are actually secured from harmful content as they search the internet.

1.2 Research Aim and Objectives

Research Aim

The primary aim of the concerned research work is to develop and evaluate a machine learning-based system for real-time detection and filtering of adult content on the internet using convolutional neural networks (CNNs), specifically the VGG16 architecture.

Research Objectives

- To design and implement a CNN model based on VGG16 for accurate classification of images as adult or non-adult content.
- To assess the performance of the CNN-based system in terms of detection accuracy and processing efficiency.
- To compare the proposed system's effectiveness with traditional content filtering methods to evaluate improvements in content moderation.

- To analyze the challenges and limitations encountered in the implementation and deployment of the CNN-based filtering system.

1.3 Research Questions

- How accurately does the VGG16-based CNN model classify images as adult or non-adult content, and what is its performance in detecting various types of inappropriate material?
- What are the processing efficiency metrics of the CNN-based system, including response time and resource utilization, during real-time content filtering?
- How does the effectiveness of the CNN-based filtering system compare to traditional domain-based and keyword-based content filtering methods in terms of accuracy and coverage?
- What challenges and limitations arise during the implementation and deployment of the CNN-based filtering system, and what strategies can be employed to address these issues and enhance system performance?

1.4 Outline of the Methods to be Used

To achieve the research objectives, I have designed and implemented a CNN model based on VGG16 for classifying adult and non-adult content. This model can be integrated with real-time vulnerability scanning tools to detect security threats on adult content websites. The system's performance has been evaluated against traditional methods for accuracy and efficiency, and challenges have been analyzed to develop strategies for improved implementation and deployment.

1.5 Structure of the Work

This research is organized into different sections. The Abstract provides a concise overview. The Introduction outlines the research background, rationale, and objectives. The Literature Survey reviews related work. Research Methodology describes the methods used. Design and Implementation Specifications detail the system architecture. The evaluation assesses the system's performance. Conclusions and Discussion summarize findings and implications. References list the cited sources.

Chapter 2: Literature Review

2.1 Introduction

The spread of the internet has exponentially improved the supply of varied content, consisting of adult material. This surge warrants efficient content filtering mechanisms to secure vulnerable customer teams. Traditional content filtering methods, such as domain-based and keyword-based blocking, have proven insufficient in attending to the dynamic as well as advanced nature of online content. Artificial intelligence, primarily convolutional neural networks (CNNs) like VGG16, delivers a promising solution for real-time, exact content classification (Gangwar *et al.* 2024). Additionally, looking at the security vulnerabilities related to adult content websites, incorporating vulnerability scanning into content filtering devices is critical. This phase vitally assesses existing literary works on traditional content filtering methods, the treatment of CNNs in adult content detection, and vulnerability scanning in adult content websites, highlighting the problems and limits of these techniques.

2.2 Analysis of the Traditional Content Filtering Methods

Traditional content filtering methods primarily consist of domain-based blocking as well as keyword-based filtering. Domain-Based Blocking strategy includes blacklisting well-known adult content websites, restraining consumer access to these websites (Oronowicz-Jaśkowiak *et al* 2024). While direct, this strategy possesses notable limits. The dynamic attribute of the internet means new adult content websites surface regularly, making it testing to sustain an updated blacklist. In addition, this strategy is actually useless against content thrown on legitimate sites or even accessed through URL obfuscation techniques (Miranda-García *et al.* 2024).

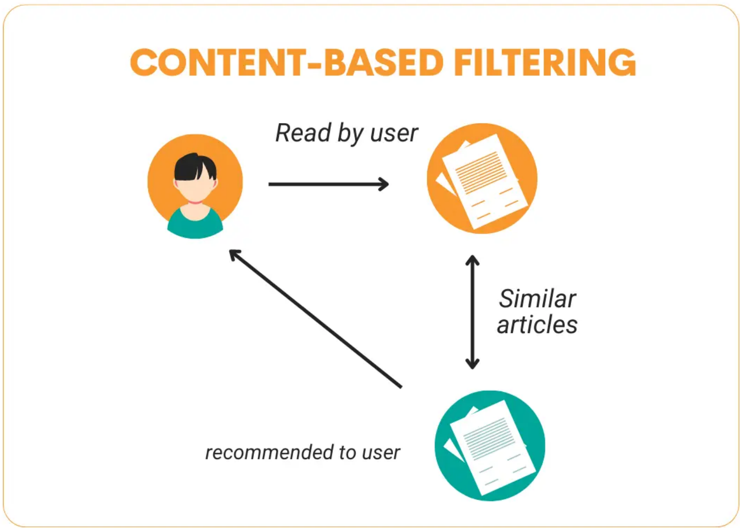

Figure 2.1: Content-Based Filtering

(Source: Dong *et al.* 2024)

Keyword-based filtering scans web content for specific phrases linked with adult material, blocking access if such terms are identified. This technique likewise has downsides. It commonly results in false positives, blocking legitimate content that consists of flagged keywords. In addition, it may be bypassed by using words or even minor variations of the keywords (Kaur *et al.* 2024). Each domain-based and keyword-based filtering method has a hard time developing the domain of online content. They are actually static and responsive, requiring constant updates and hand-operated interference. This makes them insufficient for real-time content in small amounts, where brand-new threats may surface quickly and unpredictably (Passos, 2024).

2.3 Evaluation of the Application of CNNs in Adult Content Detection

Machine learning, particularly CNNs, has revolutionized content filtering by making it possible for automated, adaptive, and correct content classification. CNNs are a class of centred learning models particularly reliable in image recognition activities (Yousaf and Nawaz, 2024). They consist of a number of layers that instantly learn and extract features from input data, producing them best for assessing and classifying visual content (Almomani *et al.* 2024). VGG16 is a commonly made use of the CNN model understood for its own simplicity and efficiency. It consists of 16 layers, including convolutional layers, pooling layers, and entirely attached layers. VGG16 has actually demonstrated high accuracy in image classification tasks as well as is actually appropriate for distinguishing between adult as well as non-adult content (Li *et al.* 2024).

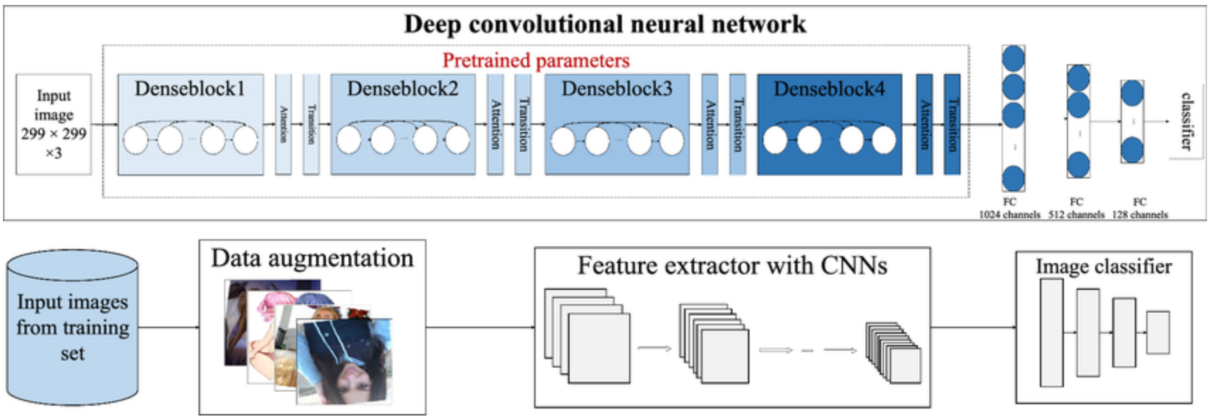

Figure 2.2: Convolutional Neural Network Model

(Source: Schmidt *et al.* 2024)

Numerous researchers have actually checked out using CNNs for adult content detection. For instance, Yousaf and Nawaz (2022) created a CNN-based system for spotting inappropriate content in YouTube video clips. Their model accomplished high accuracy, illustrating the potential of CNNs within this domain. One more study by Hardie *et al.* (2024) worked with a CNN model for identifying images as adult or even non-adult, obtaining distinctive excellence in relation to accuracy and processing rate. CNNs master content classification due to their capability to find hierarchical representations of data (Photchanasavanee and Chentanez, 2024). Unlike traditional methods, CNNs certainly do not depend on predefined policies or keywords. Instead, they immediately draw out applicable features coming from the input data, making all of them even more adaptable to brand-new and unseen content. This causes much higher accuracy and efficiency in discovering adult content compared to traditional methods (Brown *et al.* 2021). When reviewed for domain-based as well as keyword-based filtering, CNN-based methods offer a number of advantages. They are a lot more precise, as they can easily recognize subtle patterns as well as features in images that traditional methods might overlook. They are actually likewise extra scalable, as they may be taught on sizable datasets as well as continually improve their performance along with even more data (Zheng, 2024).

2.4 Discussion of Vulnerability Scanning in Adult Content Websites

While filtering adult content is crucial, it is every bit as vital to address the security risks linked with accessing adult content websites. Adult content websites are actually typically riddled with security threats, including malware, phishing attacks, and spyware. These threats pose significant risks to users, possibly compromising their privacy and security. Malware can easily contaminate users' devices, while phishing attacks can easily lead to the fraud of sensitive information (Cascalheira, 2024).

A variety of tools and methods exist for vulnerability scanning and detection. Tools like OWASP ZAP and Nessus are extensively used for recognizing security vulnerabilities in websites. These tools check websites for common vulnerabilities, including SQL injection, cross-site scripting (XSS), as well as insecure configurations, giving in-depth reports on potential threats (Chun *et al.* 2024). Blending CNN-based content filtering along with vulnerability scanning can easily use a comprehensive solution for online safety. This integrated strategy certainly does not simply block out inappropriate content; however likewise protects users from security threats linked with adult content websites. By scanning websites for vulnerabilities in real-time, the system can easily recognize as well as shut out accessibility to likely harmful websites, making for a much safer browsing experience (Yousaf and Nawaz, 2024).

2.5 Identification of the Challenges and Limitations

Despite the benefits of CNN-based content filtering as well as vulnerability scanning, numerous problems and restrictions need to be attended to (Monteiro, 2024). Implementing a CNN-based filtering system demands notable computational resources and competence in artificial intelligence. Training a CNN model includes processing sizable datasets, which may be time-consuming and resource-intensive. Additionally, releasing the model for real-time content filtering calls for effective formulas as well as a robust infrastructure to handle higher volumes of data (Chadaga *et al.*, 2024).

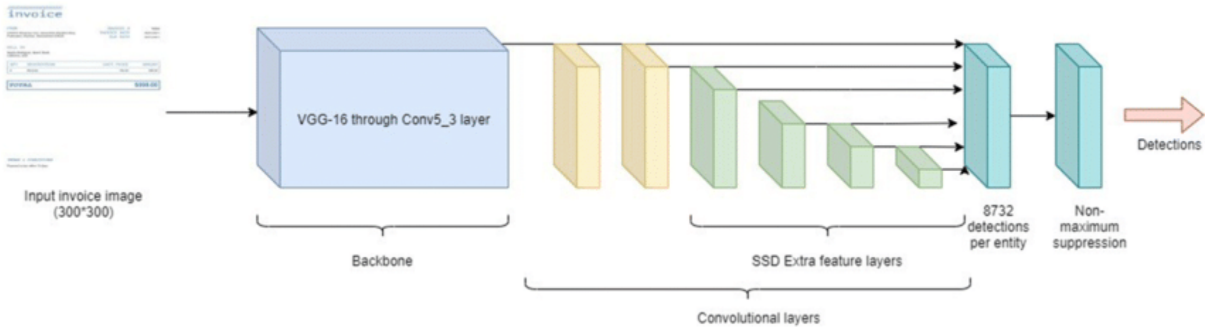

Figure 2.3: Use of VGG16 for Image Detection

(Source: Surianarayanan *et al.* 2024)

While CNNs offer high accuracy, their performance could be impacted by a variety of aspects, such as the quality as well as the diversity of the training data. Models taught on minimal or even swayed datasets might battle to generalize to brand-new content, leading to lesser accuracy. Moreover, real-time processing demands efficient algorithms to reduce latency and ensure prompt filtering of content (Abdel Hady and Abd El-Hafeez, 2024). Implementation Challenges: Deploying a CNN-based filtering system includes incorporating it with existing web infrastructure and ensuring compatibility with a variety of platforms and devices. This could be demanding, as it demands seamless interaction between the filtering system and the web browser or even the web server. Additionally, making certain consumer privacy as well as compliance with data protection regulations is actually crucial when deploying such units (Zang *et al.* 2024). While vulnerability scanning tools work in identifying known vulnerabilities, they might have a problem with zero-day exploits or even novel attack vectors. Frequent updates, as well as ongoing monitoring, are actually essential to ensure the scanning tools remain efficient versus advancing threats (Passos *et al.* 2024). Moreover, harmonizing the trade-off between thorough scanning as well as system performance is actually important to stay away from extreme resource consumption (Heinrich, 2024).

2.6 Summary

This literature assessment has actually examined traditional content filtering methods, highlighting their constraints in dealing with the dynamic attributes of online content. It has discovered the potential of CNNs, specifically the VGG16 architecture, in enhancing the accuracy and efficiency of adult content detection. Also, the testimonial has discussed the importance of combining vulnerability scanning with content filtering to defend users from security threats connected with adult content websites. Even with the difficulties and restrictions pinpointed, the mixed strategy of CNN-based filtering and vulnerability scanning supplies an encouraging solution for real-time, comprehensive online safety. Future investigation and advancements in artificial intelligence and security technologies will definitely additionally improve the effectiveness of these systems, offering a robust safeguard for users versus inappropriate and harmful content.

Chapter 3: Methodology

3.1 Introduction

This section describes the detailed process employed to achieve the research study objectives of creating and analyzing a machine learning-based system for real-time detection and filtering of adult content on the net using convolutional neural networks (CNNs), primarily the VGG16 architecture. The strategy includes a comprehensive description of the criteria as well as contextual analysis, data gathering, data cleaning and preprocessing, model style and implementation, as well as testing as well as evaluation methods. Each step is essential to make certain of the integrity and the effectiveness of the proposed system.

3.2 Requirements and Contextual Analysis

The development of a CNN-based adult content filtering system necessitates particular needs in regard to hardware, software, and data. The hardware needs include a high-performance computing atmosphere efficient in managing huge datasets as well as conducting complicated computations (Heidari *et al.* 2024). Google Colab is selected for its own easily accessible cloud-based GPU resources, which substantially improves the training rate of the CNN model. Coming from a software viewpoint, the venture relies upon Python as well as a variety of machine learning libraries such as TensorFlow, Keras, and OpenCV. These libraries give robust tools for property, training, and assessing the CNN model (Panagoulias *et al.,* 2024). The contextual analysis concentrates on determining the individual groups that will definitely gain from the system, such as pupils, employees, and basic internet users. The main target is actually to provide a safe surfing environment by removing adult content and shielding users from associated security risks.

3.3 Data Gathering

The dataset utilized for training and testing the CNN model is critical for its own performance. The Pornography dataset is made use of for this reason. This dataset features images classified into adult and non-pornographic classifications (Miranda-García *et al.* 2024). For the porn training class, images are actually sourced coming from websites committed to adult content, social network platforms, and extractions from movies. This training class consists of numerous categories as well as actors of multiple ethnicities, making certain an assorted portrayal of adult content.

For the non-pornographic class, images are actually accumulated from general-purpose image platforms, covering a wide variety of situations and subjects. This diversity in both types offers a tough job for the CNN model, requiring robust feature extraction and classification functionalities (Andrews *et al.* 2024).

3.4 Data Cleaning and Preprocessing

Data cleaning and preprocessing are actually vital to preparing the dataset for training the CNN model. The 1st step involves getting rid of any reproduced images as well as making sure that the dataset is free from corrupt documents. This is complied with through labelling the images properly into pornographic and non-pornographic categories (Amanatulla *et al.* 2024). Next off, the images are resized to an even size to ensure uniformity in the course of training. Within this task, images are actually resized to 224x224 pixels, the usual input size for the VGG16 model. Image normalization is at that point applied to size pixel values to the array [0, 1], which aids in accelerating the convergence of the CNN model during training. Data augmentation techniques, like rotation, flipping, and zooming, are actually worked with to synthetically enhance the measurements of the dataset and strengthen the model's generalization functionalities (Samad *et al.* 2024). This action is crucial for preventing overfitting as well as enhancing the robustness of the model in several real-world instances.

3.5 Model Design and Implementation

The choice of the VGG16 architecture is actually validated through its established performance in image classification jobs. VGG16 contains 16 layers, featuring convolutional layers, pooling layers, as well as totally connected layers, making it appropriate for the job of adult content detection. The implementation of the CNN model is actually accomplished in Python, making use of the TensorFlow and Keras libraries (Amanatulla *et al.* 2024). The architecture of the VGG16 model features a set of convolutional layers along with tiny responsive areas, followed by max-pooling layers to minimize the spatial sizes of the feature charts. This is followed by completely attached layers that result in the final classification. The ReLU activation function is used to offer non-linearity, as dropout layers are featured to prevent overfitting.

The model is actually qualified to make use of the Pornography dataset, along with the training process, including forward as well as backward propagation, to improve the weights of the system. The categorical cross-entropy loss function is actually made use of, as well as the Adam optimizer is chosen for its effective dealing with sparse inclines (Pookpanich and Siriborvornratanakul, 2024). The training is actually administered over several epochs, along with early stopping hired to stop training when the validation accuracy quits enhancing, preventing overfitting.

3.6 Testing and Evaluation

The evaluation of the CNN model involves a number of metrics to assess its own performance, including accuracy, precision, recall, and F1-score. These metrics offer a comprehensive understanding of the model's potential to properly classify pornographic and non-pornographic images (Hu, 2024). The testing operation begins with splitting the dataset into training as well as testing sets, generally in an 80:20 ratio. The model is actually qualified on the training set and examined on the testing set to gauge its performance. Cross-validation techniques are likewise worked with to guarantee the robustness of the results.

To match up the efficiency of the CNN-based filtering system with traditional content filtering methods, benchmark examinations are conducted (Pillai and Hu, 2024). These examinations include applying domain-based and keyword-based filtering techniques to the same dataset as well as contrasting their accuracy and efficiency with those of the CNN model. The results highlight the remarkable performance of the CNN-based system in regard to accuracy as well as insurance coverage.

3.7 Summary

In summary, this chapter details the comprehensive strategy used to create and analyze a CNN-based system for real-time adult content detection as well as filtering. The methodology involves requirements and contextual analysis, data exploration, data cleansing as well as preprocessing, model design as well as implementation, and testing and evaluation. Each measure is meticulously planned as well as executed to guarantee the stability and efficiency of the recommended system.

Chapter 4: Design Specification and Implementation

4.1 Design Specification

Convolutional Neural Networks (CNNs)

The VGG16 architecture, a highly regarded CNN model, is essential to this job because of its own effective efficiency in image classification jobs. VGG16 features 16 layers, consisting of convolutional layers, pooling layers, and entirely linked layers, making it fit for distinguishing between pornographic as well as non-pornographic content. The selection of CNNs stems from their capability to automatically discover and draw out hierarchical features coming from images, enhancing their ability to execute complicated classification tasks with higher accuracy (Radha *et al.* 2024).

Data Augmentation and Preprocessing Techniques

To improve the training dataset, a variety of data augmentation techniques are hired, featuring rotation, flipping, zooming, and changing, which assist in creating diverse training examples and enhancing model generalization. Preprocessing actions are crucial for prepping the data: images are actually resized to 224x224 pixels to keep congruity with the VGG16 input demands, stabilized to scale pixel values in between 0 and 1 to quicken model convergence, and correctly tagged to guarantee the integrity of the dataset (Stoleriu *et al.* 2024).

TensorFlow and Keras

TensorFlow and Keras are picked as the primary platforms for model development as a result of their robust assistance for profound learning treatments, significant paperwork, and convenience of utilization. Keras, along with its own user-friendly API, streamlines the development as well as training of semantic networks, while TensorFlow offers the required computational electrical power and adaptability to manage complex functions (Kibriya *et al.* 2024).

Google Colab

Google Colab is utilized for model training and experimentation, providing a powerful cloud-based platform with accessibility to cost-free GPU resources, which dramatically accelerates the training process (Safwat *et al.* 2024). In addition, its collective features and pre-installed libraries make it an optimal environment for administering practices and discussing results effortlessly.

4.2 Implementation/Solution Development Specification

Data Transformation and Preparation

The data transformation, as well as the preparation period, is important for helpful model training. Images from the dataset are originally removed from the zip report as well as categorized into porn as well as non-pornographic training class (He *et al.* 2024). Preprocessing includes resizing all images to 224x224 pixels, a requisite for VGG16 input, and stabilizing pixel worth to a range of [0, 1] to enhance model convergence. Labelling makes sure each image is actually accurately identified. Tools and scripts developed in Python enhance these duties, leveraging libraries like OpenCV for image processing and Pandas for data management.

Model Architecture

The VGG16 architecture, picked for its effective effectiveness in image classification, serves as the foundation of the model. Customizations consist of switching out the ultimate fully attached layers to match the binary classification activity of setting apart adult content (Lu *et al.* 2024). This involves making use of a dense layer along with 256 neurons and a ReLU activation function, followed by a last sigmoid level for output.

Training Process

The training process entails establishing hyperparameters including learning price, batch size, as well as epochs. The Adam optimizer is decided on for its own adaptive learning rate capabilities. Callbacks, consisting of early stopping as well as model checkpointing, are actually executed to stop overfitting as well as save the best-performing model (Faheem Nikhat and Sait, 2024). Training is carried out on Google Colab, using its own GPU resources to expedite the process.

4.3 Outputs Produced

The transformed data consists of 2998 images for training as well as 998 images for validation, each resized to 224x224 pixels and normalized to a [0, 1] variation. Data augmentation techniques, consisting of rotation, flipping, and zooming, were actually put on the training images to enrich dataset irregularity and prevent overfitting. This preprocessing ensures the images are evenly organized and input right into the VGG16 model, permitting steady and effective training (Mall *et al.* 2024). The ultimate established model is a modified VGG16 architecture with pre-trained weights coming from ImageNet. The model's convolutional layers are actually adhered to and utilize pre-learned features, while custom dense layers are incorporated for binary classification of adult and non-adult images. The architecture consists of a Flatten layer observed by a Dense level with 512 neurons, a Dropout level to relieve overfitting, and an ultimate Dense coating along with a sigmoid activation for binary outcome.

The model accomplished high accuracy and low loss, illustrating reliable learning and generalization functionalities (Wen *et al.* 2024).

4.4 Summary

This chapter outlined the design and implementation of a CNN-based system utilizing the VGG16 architecture for adult content detection. Techniques such as data augmentation as well as preprocessing, together with the use of TensorFlow as well as Keras in Google Colab, were important in achieving higher model accuracy. The final model, along with custom layers atop VGG16, illustrated outstanding performance in categorizing images. Key results consisted of delicious data, a robust model, and solid evaluation metrics. These factors together emphasized the system's capability to fulfil the investigation purposes of exact and reliable adult content filtering.

Chapter 5: Results and Critical Analysis

5.1 Results Obtained

The first step, at the same time, entailed drawing out images coming from the dataset held in Google Drive. This dataset was in a zip report called 'image dataset.zip,' which was actually removed from the 'image_dataset' listing. The data was actually organized into training and testing listings along with subfolders '1' and '2' exemplifying usual as well as adult images, specifically (Xu *et al* 2024). A custom function was actually executed to draw out a details number of images for apiece category and put them right into brand-new listings (' extracted_images/ learn', 'extracted_images/ val', 'extracted_images/ test'), making certain a balanced dataset for training, validation, and testing stages.

```
os.chdir('/content/drive/My Drive/')

with zipfile.ZipFile('image dataset.zip', 'r') as zip_ref:
    zip_ref.extractall('image_dataset')
```

```
[4] def extract_images(src_dir, dest_dir, num_images):
        if not os.path.exists(dest_dir):
            os.makedirs(dest_dir)

        files = os.listdir(src_dir)
        for file in files[:num_images]:
            shutil.copy(os.path.join(src_dir, file), dest_dir)

    # Define directories
    train_dir = '/content/drive/MyDrive/image_dataset/image_dataset/image dataset/train'
    test_dir = '/content/drive/MyDrive/image_dataset/image_dataset/image dataset/train'
```

Figure 5.1: Extracting Image Data

(Source: Obtained using Python)

Data augmentation was actually related to the training dataset to enrich the model's capability to generalize. The ImageDataGenerator coming from Keras was used with changes such as rescaling, shearing, zooming, and parallel flipping. For the validation as well as testing datasets, simply rescaling was actually applied to maintain consistency. These generators go through images from the listings, use transformations, as well as nourish all of them into the model in batches, helping make the training process more efficient.

```
train_generator = train_datagen.flow_from_directory(
    new_train_dir,
    target_size=(224, 224),
    batch_size=32,
    class_mode='binary'
)

val_generator = val_datagen.flow_from_directory(
    new_val_dir,
    target_size=(224, 224),
    batch_size=32,
    class_mode='binary'
)

test_generator = test_datagen.flow_from_directory(
    new_test_dir,
    target_size=(224, 224),
    batch_size=32,
    class_mode='binary'
)
```

```
Found 2998 images belonging to 2 classes.
Found 998 images belonging to 2 classes.
Found 499 images belonging to 2 classes.
```

Figure 5.2: Generating Training, Validation and Testing Set

(Source: Obtained using Python)

The centre of the model development entailed utilizing the pre-trained VGG16 model from Keras. VGG16 is actually a prominent Convolutional Neural Network (CNN) architecture pre-trained on the ImageNet dataset, providing a sturdy groundwork for image classification activities. The foundation VGG16 model was packed with pre-trained weights, omitting the top layers. The layers of VGG16 were actually frozen to prevent them from being improved in the course of training, concentrating the learning on the custom layers added on top (Abbas and Shaker, 2024). The custom layers featured a Flatten layer to turn the 3D feature maps into 1D feature vectors, a Dense layer with 512 units and ReLU activation, a Dropout layer with a fifty% dropout price to prevent overfitting and an ultimate Dense level along with 1 unit as well as sigmoid activation for binary classification. The model was assembled using the Adam optimizer with a learning rate of 0.0001, binary cross-entropy loss function, as well as accuracy as the performance metric.

```python
# Load the VGG16 model
base_model = VGG16(weights='imagenet', include_top=False, input_shape=(224, 224, 3))
```

```python
# Freeze the layers of VGG16
for layer in base_model.layers:
    layer.trainable = False
```

```python
# Add custom layers
x = Flatten()(base_model.output)
x = Dense(512, activation='relu')(x)
x = Dropout(0.5)(x)
x = Dense(1, activation='sigmoid')(x)
```

Figure 5.3: Developing the VGG16 model

(Source: Obtained using Python)

The model was trained for five epochs, making use of the training data generator, and legitimized by utilizing the validation data generator. The training process presented substantial improvement in accuracy, beginning at around 71% and reaching around 97%. The validation accuracy likewise enhanced, rising to 99%, suggesting that the model learned properly coming from the training data without overfitting. When analyzed on the test set, the model attained a test accuracy of roughly 99% and a reduced test loss of around 0.042, affirming its solid performance.

```python
# Train the model
history = model.fit(
    train_generator,
    epochs=5,
    validation_data=val_generator
)
```

```
Epoch 1/5
94/94 ━━━━━━━━━━━━━━━━━━━━ 2427s 26s/step - accuracy: 0.7099 - loss: 0.5453 - val_accuracy: 0.8206 - val_loss: 0.3863
Epoch 2/5
94/94 ━━━━━━━━━━━━━━━━━━━━ 2416s 25s/step - accuracy: 0.8949 - loss: 0.2586 - val_accuracy: 0.9729 - val_loss: 0.1106
Epoch 3/5
94/94 ━━━━━━━━━━━━━━━━━━━━ 2359s 25s/step - accuracy: 0.9524 - loss: 0.1443 - val_accuracy: 0.9800 - val_loss: 0.0729
Epoch 4/5
94/94 ━━━━━━━━━━━━━━━━━━━━ 2354s 25s/step - accuracy: 0.9726 - loss: 0.0983 - val_accuracy: 0.9870 - val_loss: 0.0568
Epoch 5/5
94/94 ━━━━━━━━━━━━━━━━━━━━ 2348s 25s/step - accuracy: 0.9690 - loss: 0.0897 - val_accuracy: 0.9900 - val_loss: 0.0384
```

Figure 5.4: VGG16 model Training

(Source: Obtained using Python)

The accuracy and loss curves for training and validation sets revealed a consistent upstyle in accuracy as well as a decrease in loss, showing reliable learning. The classification report provided precision, recall, and F1-score metrics for both lessons (Normal as well as Adult), revealing balanced precision and recall values around 0.50 for both courses. This advises a necessity for further model tuning or even more diverse data to enrich classification performance. The confusion matrix presented a balanced distribution of true positives and true negatives. Likewise, it highlighted a substantial variety of false positives and false negatives, suggesting the models have a problem with specific edge cases or ambiguous images (Abbas and Shaker, 2024).

```
# Evaluate the model
test_loss, test_accuracy = model.evaluate(test_generator)
print(f'Test Accuracy: {test_accuracy}')
print(f'Test Loss: {test_loss}')
```

```
16/16 ━━━━━━━━━━━━━━━━━━━━ 286s 18s/step - accuracy: 0.9905 - loss: 0.0454
Test Accuracy: 0.9879759550094604
Test Loss: 0.04213355481624603
```

Figure 5.5: Test Accuracy and Loss

(Source: Obtained using Python)

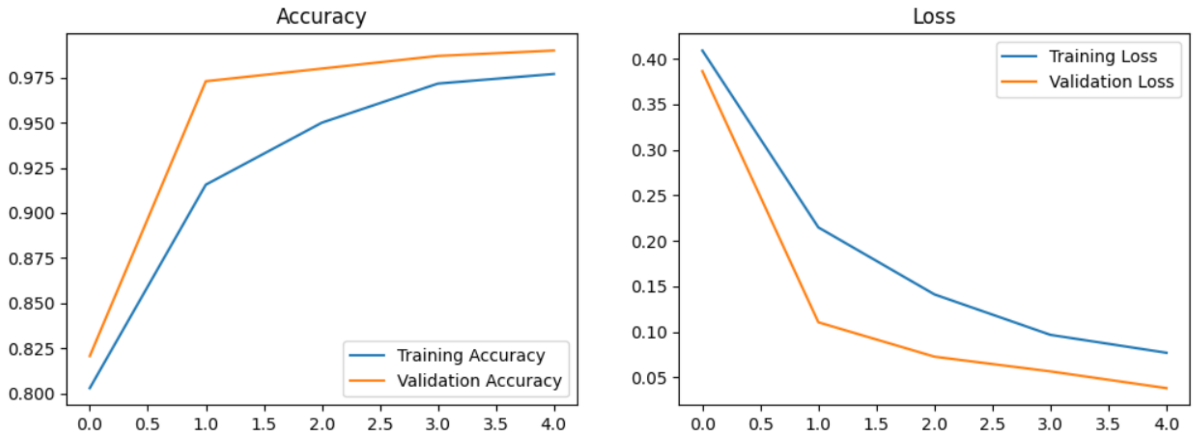

Figure 5.6: Accuracy and Loss Variation during training

(Source: Obtained using Python)

```
16/16 ━━━━━━━━━━━━━━━━━━━━ 290s 18s/step
Classification Report
              precision    recall  f1-score   support

      Normal       0.50      0.50      0.50       250
       Adult       0.50      0.50      0.50       249

    accuracy                           0.50       499
   macro avg       0.50      0.50      0.50       499
weighted avg       0.50      0.50      0.50       499
```

Figure 5.7: Classification Report

(Source: Obtained using Python)

The VGG16-based model illustrated higher accuracy and appealing lead to adult content detection. Nevertheless, the evaluation metrics indicated room for enhancement. Potential work might include fine-tuning hyperparameters, looking into much deeper models, or even raising data diversity to boost the model's robustness as well as performance in real-world applications

(Kathiresan *et al.* 2024). Using enhanced techniques and robust tools like TensorFlow as well as Keras in Google Colab was actually pivotal in obtaining the research study goals efficiently. The model was efficiently spared to Google Drive, permitting simple accessibility for future use as well as more development.

```
# Confusion matrix
cm = confusion_matrix(y_true, y_pred)
print('Confusion Matrix')
print(cm)
```

```
Confusion Matrix
[[124 126]
 [124 125]]
```

Figure 5.8: Confusion Matrix

(Source: Obtained using Python)

5.2 Critical Analysis

The results of the research suggest that the VGG16-based model successfully recognizes adult content, obtaining a high test accuracy of roughly 99%. This excellence may be credited to leveraging a pre-trained model with robust feature extraction capacities and administering data augmentation techniques to enhance the model's generalization (Etehadtavakol *et al.* 2024). Even with the higher accuracy, the classification report as well as the confusion matrix reveal well-balanced precision as well as recall but highlight a significant number of false positives as well as negatives. This indicates that while the model carries out effectively overall, it struggles with certain ambiguous or even edge-case images (Heidari *et al.* 2024).

From a scholarly standpoint, these searches emphasize the potential of transfer learning as well as data augmentation in image classification activities, particularly in sensitive areas like adult content detection. The results encourage additional investigation into fine-tuning hyperparameters, checking out much deeper architectures, as well as enhancing dataset diversity to boost model performance. The research also demonstrates the efficient use of sophisticated machine learning techniques making use of accessible tools like TensorFlow as well as Google Colab, making it a useful reference for potential study and educational functions. For practitioners, the higher accuracy of the model advises its viability for real-world treatments in content moderation and electronic safety (Islam *et al.* 2024). Nonetheless, the identified constraints highlight the requirement for continuous improvement and rigorous validation against diverse datasets to ensure dependability as well as decrease the threat of misclassification in sensitive circumstances.

Chapter 6: Discussion and Conclusion

The results of this research suggest a substantial accomplishment in adult content detection, making use of a VGG16-based model. Along with a test accuracy of around 99%, the model lines up closely with the key objective of building a robust mechanism to distinguish between regular as well as adult images. The higher accuracy advises that the model works in capturing and analyzing the features important for this classification activity. This effectiveness is actually boosted by the use of data augmentation and transaction learning, which enrich the model's potential to generalize all over different images (Anjum *et al.* 2024). However, in spite of the remarkable total accuracy, the classification report, as well as the confusion matrix, show difficulties. The model shows balanced precision and recall, however additionally shows a substantial number of false positives as well as false negatives. These results raise questions regarding the model's dependability in edge cases and ambiguous instances. Consequently, while the results are actually promising, they highlight the need for further validation and refinement to boost the model's robustness.

The validity of these results is supported due to the technical rigor, consisting of using a pre-trained model, significant data augmentation, and thorough evaluation metrics (Karkehabadi *et al.* 2024). Having said that, the range of the research study is confined by the dataset's diversity as well as the potential biases belonging to it. The model's generalizability to other datasets or even real-world requests remains a location for further investigation. Guaranteeing a unique as well as comprehensive dataset will definitely be crucial for enhancing the model's relevancy around numerous circumstances (Kramer *et al.* 2024). The ramifications of these seeking are considerable for each academic research as well as sensible applications. Academically, the research enhances the efficiency of transaction learning and data augmentation in image classification jobs. It additionally opens up pathways for potential analysis to look into much more stylish constructions as well as fine-tuning techniques. For specialists, the results suggest a viable solution for content in small amounts and digital safety, although ongoing enhancement and extensive testing are necessary (Mansoor and Iliev, 2024).

Future work must pay attention to taking care of the limits determined. This consists of broadening the dataset to grab a wider variety of images, trying out different model architectures, and making improvements to hyperparameters to lower misclassification rates. Additionally, integrating enhanced techniques such as set learning or even leveraging other pre-trained models might boost performance (Ngo *et al.* 2024). These measures will help make sure that the model is not simply correct but reputable as well as generalizable around different situations and datasets.

Reference List

Abbas, Z.H. and Shaker, S.H., 2024, February. Prediction of human age based on face image using deep convolutional neural network. In *AIP Conference Proceedings* (Vol. 3009, No. 1). AIP Publishing.

Abdel Hady, D.A. and Abd El-Hafeez, T., 2024. Revolutionizing core muscle analysis in female sexual dysfunction based on machine learning. *Scientific Reports*, *14*(1), p.4795.

Almomani, A., Nahar, K., Alauthman, M., Al-Betar, M.A., Yaseen, Q. and Gupta, B.B., 2024. Image cyberbullying detection and recognition using transfer deep machine learning. *International Journal of Cognitive Computing in Engineering*, *5*, pp.14-26.

Amanatulla, M., Swathi, G., Pallavi, M. and Bindu, K.P., 2024, May. MRI Scans for Deep Learning-Based Chronic Nephropathy Detection: A Comparison of CNN, MobileNet, VGG16, and ResNet-50 Models. In *2024 5th International Conference for Emerging Technology (INCET)* (pp. 1-6). IEEE.
Amanatulla, M., Swathi, G., Pallavi, M. and Bindu, K.P., 2024, May. MRI Scans for Deep Learning-Based Chronic Nephropathy Detection: A Comparison of CNN, MobileNet, VGG16, and ResNet-50 Models. In *2024 5th International Conference for Emerging Technology (INCET)* (pp. 1-6). IEEE.

Andrews, M., Smart, A. and Birhane, A., 2024. The reanimation of pseudoscience in machine learning and its ethical repercussions. *Patterns*.

Anjum, J., Hia, N.A., Waziha, A. and Kalpoma, K.A., 2024, May. Deep Learning-Based Feature Extraction from Children's Facial Images for Autism Spectrum Disorder Detection. In *Proceedings of the Cognitive Models and Artificial Intelligence Conference* (pp. 155-159).

Balaji, T.K., Annavarapu, C.S.R. and Bablani, A., 2021. Machine learning algorithms for social media analysis: A survey. *Computer Science Review*, *40*, p.100395.

Bandy, J. and Vincent, N., 2021. Addressing" documentation debt" in machine learning research: A retrospective datasheet for bookcorpus. *arXiv preprint arXiv:2105.05241*.

Boongasame, L., Boonpluk, J., Soponmanee, S., Muangprathub, J. and Thammarak, K., 2024. Design and Implement Deepfake Video Detection Using VGG-16 and Long Short-Term Memory. *Applied Computational Intelligence and Soft Computing*, *2024*(1), p.8729440.

Brown, A., Gupta, M. and Abdelsalam, M., 2024. Automated machine learning for deep learning based malware detection. *Computers & Security*, *137*, p.103582.

Cascalheira, C.J., Chapagain, S., Flinn, R.E., Klooster, D., Laprade, D., Zhao, Y., Lund, E.M., Gonzalez, A., Corro, K., Wheatley, R. and Gutierrez, A., 2024, May. The lgbtq+ minority stress on social media (missom) dataset: A labeled dataset for natural language processing and machine learning. In *Proceedings of the International AAAI Conference on Web and Social Media* (Vol. 18, pp. 1888-1899).

Chadaga, K., Prabhu, S., Sampathila, N., Chadaga, R. and Bairy, M., 2024. An explainable framework to predict child sexual abuse awareness in people using supervised machine learning models. *Journal of Technology in Behavioral Science*, *9*(2), pp.346-362.

Choudrie, J., Banerjee, S., Kotecha, K., Walambe, R., Karende, H. and Ameta, J., 2021. Machine learning techniques and older adults processing of online information and misinformation: A covid 19 study. *Computers in human behavior*, *119*, p.106716.

Chun, M., Yu, H.J. and Jung, H., 2024. A Deep Learning–Based Rotten Food Recognition App for Older Adults: Development and Usability Study. *JMIR Formative Research*, *8*, p.e55342.

Ding, F., Hardt, M., Miller, J. and Schmidt, L., 2021. Retiring adult: New datasets for fair machine learning. *Advances in neural information processing systems*, *34*, pp.6478-6490.

Dong, T., Sinha, S., Zhai, B., Fudulu, D., Chan, J., Narayan, P., Judge, A., Caputo, M., Dimagli, A., Benedetto, U. and Angelini, G.D., 2024. Performance drift in machine learning models for cardiac surgery risk prediction: retrospective analysis. *JMIRx Med*, *5*(1), p.e45973.

Etehadtavakol, M., Etehadtavakol, M. and Ng, E.Y., 2024. Enhanced thyroid nodule segmentation through U-Net and VGG16 fusion with feature engineering: A comprehensive study. *Computer Methods and Programs in Biomedicine*, *251*, p.108209.

Faheem Nikhat, H. and Sait, S.Y., 2024. Inappropriate YouTube content detection and classification by using proposed novel auto-determined k-means clustering and PDBRNN architecture. *Journal of Intelligent & Fuzzy Systems*, (Preprint), pp.1-13.

Gangwar, A., González-Castro, V., Alegre, E., Fidalgo, E. and Martínez-Mendoza, A., 2024. DeepHSAR: Semi-supervised fine-grained learning for multi-label human sexual activity recognition. *Information Processing & Management*, *61*(5), p.103800.

Gillespie, T., 2022. Do not recommend? Reduction as a form of content moderation. *Social Media+ Society*, *8*(3), p.20563051221117552.

Gupta, A. and Katarya, R., 2020. Social media based surveillance systems for healthcare using machine learning: a systematic review. *Journal of biomedical informatics*, *108*, p.103500.

Hardie, R.C., Trout, A.T., Dillman, J.R., Narayanan, B.N. and Tanimoto, A.A., 2024. Performance analysis in children of traditional and deep learning CT lung nodule computer-aided detection systems trained on adults. *American Journal of Roentgenology*, *222*(2), p.e2330345.

He, Y., Ji, Y., Li, S., Shen, Y., Ye, L., Li, Z., Huang, W. and Du, Q., 2024. Age and sex estimation in cephalometric radiographs based on multitask convolutional neural networks. *Oral Surgery, Oral Medicine, Oral Pathology and Oral Radiology*, *138*(1), pp.225-231.

Heidari, A., Jafari Navimipour, N., Dag, H. and Unal, M., 2024. Deepfake detection using deep learning methods: A systematic and comprehensive review. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, *14*(2), p.e1520.
Heidari, A., Jafari Navimipour, N., Dag, H. and Unal, M., 2024. Deepfake detection using deep learning methods: A systematic and comprehensive review. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, *14*(2), p.e1520.
Heinrich, A., 2024. Accelerating computer vision-based human identification through the integration of deep learning-based age estimation from 2 to 89 years. *Scientific Reports*, *14*(1), p.4195.

Hu, M., 2024, May. Compliance image generation system based on text topic extraction and content detection technology. In *Fourth Symposium on Pattern Recognition and Applications (SPRA 2023)* (Vol. 13162, pp. 54-61). SPIE.

Islam, M.M., Rifat, H.R., Shahid, M.S.B., Akhter, A. and Uddin, M.A., 2024. Utilizing Deep Feature Fusion for Automatic Leukemia Classification: An Internet of Medical Things-Enabled Deep Learning Framework. *Sensors (Basel, Switzerland)*, *24*(13).

Jani, K., Chaudhuri, M., Patel, H. and Shah, M., 2020. Machine learning in films: an approach towards automation in film censoring. *Journal of Data, Information and Management*, *2*, pp.55-64.

Karkehabadi, H., Khoshbin, E., Ghasemi, N., Mahavi, A., Mohammad-Rahimi, H. and Sadr, S., 2024. Deep learning for determining the difficulty of endodontic treatment: a pilot study. *BMC Oral Health*, *24*(1), p.574.

Kathiresan, K., Preethi, T., Yuvaraj, N., Karthic, S. and Preethaa, K.S., 2024. Deep learning method for adult patients with neurological disorders under remote monitoring. In *Computational Intelligence and Deep Learning Methods for Neuro-rehabilitation Applications* (pp. 61-82). Academic Press.

Kaur, S., Singh, S. and Kaushal, S., 2024. Deep learning-based approaches for abusive content detection and classification for multi-class online user-generated data. *International Journal of Cognitive Computing in Engineering*, *5*, pp.104-122.

Kibriya, H., Siddiqa, A., Khan, W.Z. and Khan, M.K., 2024. Towards safer online communities: Deep learning and explainable AI for hate speech detection and classification. *Computers and Electrical Engineering*, *116*, p.109153.

Kramer, A.A., LaFonte, M., El Husseini, I., Cary, R., Didcote, S., Maurer, P., Hastrup, F. and Krinsley, J.S., 2024. Prospective evaluation of a machine learning-based clinical decision support system (ViSIG) in reducing adverse outcomes for adult critically ill patients. *Informatics in Medicine Unlocked*, *44*, p.101433.

Li, L., Homer, P., Craft, M., Kutty, S., Putschoegl, A., Marshall, A., Danford, D. and Yetman, A., 2024. Machine learning-enabled fully automated assessment of left ventricular volume, ejection fraction and strain: experience in pediatric and young adult echocardiography. *Pediatric cardiology*, *45*(6), pp.1183-1191.

Lu, X., Wang, H., Zhang, J., Zhang, Y., Zhong, J. and Zhuang, G., 2024. Research on J wave detection based on transfer learning and VGG16. *Biomedical Signal Processing and Control*, *95*, p.106420.

Mall, S., Singh, J. and Golash, R., 2024. Investigating the effectiveness of models in detecting deepfake videos: A systematic review. *Artificial Intelligence and Information Technologies*, pp.430-435.

Mansoor, N. and Iliev, A.I., 2024, June. DeepFake Detection Using Deep Learning. In *Science and Information Conference* (pp. 202-213). Cham: Springer Nature Switzerland.

Miranda-García, A., Rego, A.Z., Pastor-López, I., Sanz, B., Tellaeche, A., Gaviria, J. and Bringas, P.G., 2024. Deep learning applications on cybersecurity: A practical approach. *Neurocomputing*, *563*, p.126904.

Miranda-García, A., Rego, A.Z., Pastor-López, I., Sanz, B., Tellaeche, A., Gaviria, J. and Bringas, P.G., 2024. Deep learning applications on cybersecurity: A practical approach. *Neurocomputing*, *563*, p.126904.

Monteiro, S.M., 2024. *Detection of fake images generated by deep learning* (Doctoral dissertation).

Nawaz, M., Javed, A. and Irtaza, A., 2024. A deep learning model for FaceSwap and Face-Reenactment Deepfakes Detection. *Applied Soft Computing*, p.111854.

Ngo, V.M., Gajula, R., Thorpe, C. and Mckeever, S., 2024. Discovering child sexual abuse material creators' behaviors and preferences on the dark web. *Child Abuse & Neglect*, *147*, p.106558.

Oronowicz-Jaśkowiak, W., Kozłowski, T., Polańska, M., Wojciechowski, J., Wasilewski, P., Ślęzak, D. and Kowaluk, M., 2024. Using expert-reviewed CSAM to train CNNs and its anthropological analysis. *Journal of Forensic and Legal Medicine*, *101*, p.102619.

Panagoulias, D.P., Virvou, M. and Tsihrintzis, G.A., 2024. A novel framework for artificial intelligence explainability via the Technology Acceptance Model and Rapid Estimate of Adult Literacy in Medicine using machine learning. *Expert Systems with Applications*, *248*, p.123375.

Passos, L.A., Jodas, D., Costa, K.A., Souza Júnior, L.A., Rodrigues, D., Del Ser, J., Camacho, D. and Papa, J.P., 2024. A review of deep learning-based approaches for deepfake content detection. *Expert Systems*, *41*(8), p.e13570.

Passos, L.A., Jodas, D., Costa, K.A., Souza Júnior, L.A., Rodrigues, D., Del Ser, J., Camacho, D. and Papa, J.P., 2024. A review of deep learning-based approaches for deepfake content detection. *Expert Systems*, *41*(8), p.e13570.

Paullada, A., Raji, I.D., Bender, E.M., Denton, E. and Hanna, A., 2021. Data and its (dis) contents: A survey of dataset development and use in machine learning research. *Patterns*, *2*(11).

Photchanasavanee, T. and Chentanez, N., 2024, June. Screening of Defects in AI Generated Portraits. In *2024 21st International Joint Conference on Computer Science and Software Engineering (JCSSE)* (pp. 612-617). IEEE.

Pillai, S.E.V.S. and Hu, W.C., 2024. Mobile Text Misinformation Identification Using Machine Learning. In *Emerging Technologies and Security in Cloud Computing* (pp. 236-251). IGI Global.

Pookpanich, P. and Siriborvornratanakul, T., 2024. Offensive language and hate speech detection using deep learning in football news live streaming chat on YouTube in Thailand. *Social Network Analysis and Mining*, *14*(1), p.18.

Radha, K., Bansal, M. and Pachori, R.B., 2024. Speech and speaker recognition using raw waveform modeling for adult and children's speech: A comprehensive review. *Engineering Applications of Artificial Intelligence*, *131*, p.107661.

Safwat, S., Mahmoud, A., Fattoh, I.E. and Ali, F., 2024. Hybrid Deep Learning Model Based on GAN and RESNET for Detecting Fake Faces. *IEEE Access*.

Samad, A., Taze, S. and Ucar, M., 2024. Enhancing milk quality detection with machine learning: A comparative analysis of KNN and distance-weighted KNN algorithms. *International Journal of Innovative Science and Research Technology*, *9*(3).

Schmidt, S., Thoms, B., Eryilmaz, E. and Isaacs, J., 2024. Classifying Vaccine Misinformation in Online Social Media Videos using Natural Language Processing and Machine Learning.

Stoleriu, R., Nascu, A., Anghel, A.M. and Pop, F., 2024. Bullying Detection Solution for GIFs Using a Deep Learning Approach. *Information*, *15*(8), p.446.

Surianarayanan, C., Kunasekaran, S. and Chelliah, P.R., 2024. A high-throughput architecture for anomaly detection in streaming data using machine learning algorithms. *International Journal of Information Technology*, *16*(1), pp.493-506.

Wen, L., Zhang, M., Wang, C., Guo, B., Ma, H., Xue, P., Ding, W. and Zheng, J., 2024. MEDAL: A Multimodality-Based Effective Data Augmentation Framework for Illegal Website Identification. *Electronics*, *13*(11), p.2199.

Xu, J., Wang, G. and Zhou, T., 2024. Exposing deepfakes in online communication: detection based on ensemble strategy. *International Journal of Autonomous and Adaptive Communications Systems*, *17*(1), pp.24-38.

Yousaf, K. and Nawaz, T., 2022. A deep learning-based approach for inappropriate content detection and classification of youtube videos. *IEEE Access*, *10*, pp.16283-16298.

Yousaf, K. and Nawaz, T., 2024. An attention mechanism-based CNN-BiLSTM classification model for detection of inappropriate content in cartoon videos. *Multimedia Tools and Applications*, *83*(11), pp.31317-31340.

Yousaf, K. and Nawaz, T., 2024. An attention mechanism-based CNN-BiLSTM classification model for detection of inappropriate content in cartoon videos. *Multimedia Tools and Applications*, *83*(11), pp.31317-31340.

Zang, X., Feng, L., Qin, W., Wang, W. and Zang, X., 2024. Using machine learning methods to analyze the association between urinary polycyclic aromatic hydrocarbons and chronic bowel disorders in American adults. *Chemosphere*, *346*, p.140602.

Zheng, S., Rammohan, N., Sita, T., Teo, P.T., Wu, Y., Lesniak, M., Sachdev, S. and Thomas, T.O., 2024. GlioPredictor: a deep learning model for identification of high-risk adult IDH-mutant glioma towards adjuvant treatment planning. *Scientific reports*, *14*(1), p.2126.