ISSN: 2959-6386 (Online), Volume 2, Issue 2

Journal of Knowledge Learning and Science Technology

Journal homepage: https://jklst.org/index.php/home

Advancing NSFW Detection in AI: Training Models to Detect Drawings, Animations, and Assess Degrees of Sexiness

Nicolas Guzman
Universidad de La Sabana, Colombia

Abstract

This research explores the advancement of NSFW (Not Safe for Work) detection in AI by training models to detect NSFW content in drawings, animations, and assess degrees of sexiness. Leveraging the NSFWJS library as a foundation, we conduct a comprehensive investigation into enhancing the capabilities of existing NSFW detection models. Through a systematic approach encompassing data collection, annotation, model training, and evaluation, we fine-tune the NSFWJS model to effectively identify NSFW content across diverse media types. Our research addresses the growing need for robust NSFW detection in AI applications, particularly in scenarios involving non-photographic content and nuanced assessments of sexual content. By expanding the capabilities of NSFW detection models, this work contributes to creating safer online environments and enabling more responsible content moderation practices.

Keywords: NSFW detection, AI, NSFWJS library, drawings, animations, sexiness assessment, model training

Article Information:
Article history: *Received:01/10/2023Accepted:3/10/2023* *Online: 12/10/2023 Published: 12/10/2023*
DOI: *https://doi.org/10.60087/jklst.vol2.n2.p294*
[i] ***Correspondence author: Nicolas Guzman***

Introduction

Web archiving (WA) is dedicated to preserving segments of the World Wide Web (WWW) for future access. NSFW JS, a prominent WA initiative, houses a vast repository of web content, including image files. However, within this repository, there exist images containing nudity and pornography, deemed Not Suitable For Work

(NSFW) and potentially offensive to users. This study proposes a solution to classify NSFW images identified within NSFW JS using deep neural network methodologies.

To facilitate NSFW image classification, a sizable dataset of images is compiled using NSFW JS data. Subsequently, two pre-trained neural network models, ResNet and SqueezeNet, are assessed and refined for the NSFW classification task leveraging the dataset. Evaluation of these models initially yielded accuracy scores of 93% and 72%, respectively. Following a fine-tuning stage, the accuracy of these models was enhanced to 94% and 89%, respectively.

The developed solution is seamlessly integrated into the Arquivo.pt Image Search System, accessible at https://nsfwjs.com/. This integration ensures that users can efficiently identify and filter out NSFW content while exploring the extensive image repository maintained by NSFW JS, thereby enhancing the overall user experience and usability of the platform.

The proliferation of digital content on the internet has brought about the need for effective methods to detect and moderate potentially inappropriate or NSFW (Not Safe for Work) material. With the advent of AI and machine learning, automated NSFW detection systems have become increasingly prevalent, aiding in content moderation across various online platforms. However, while existing models excel at identifying NSFW content in photographs, they often struggle with non-photographic media types such as drawings, animations, and assessing degrees of sexual content or "sexiness." In response to these challenges, this research endeavors to advance NSFW detection in AI by training models to detect NSFW content in drawings, animations, and accurately assess degrees of sexiness.

Building upon the NSFWJS library, a widely-used tool for NSFW detection in images, we embark on a comprehensive exploration to enhance its capabilities. Our research seeks to address the limitations of current NSFW detection models by extending their proficiency to encompass diverse media types beyond photographs. By focusing on drawings and animations, we aim to develop models capable of identifying NSFW content across a broader spectrum of visual media, thereby improving the effectiveness of content moderation efforts in online environments.

Moreover, we delve into the nuanced aspect of sexiness assessment, recognizing the importance of not only detecting NSFW content but also gauging its level of sexual suggestiveness. Through meticulous data collection, annotation, and model training, we endeavor to equip AI systems with the ability to discern varying degrees of sexual content, catering to the multifaceted nature of online content moderation requirements.

By advancing NSFW detection capabilities in AI, this research contributes to fostering safer online environments, protecting users from encountering inappropriate content, and facilitating more nuanced content moderation practices. Furthermore, our work holds implications for a wide range of applications, including social media platforms, content-sharing websites, and online communities, where effective NSFW detection is paramount for maintaining a conducive and respectful online environment.

Objectives

Objective 1: Enhance NSFW Detection in Drawings and Animations

- Develop techniques to preprocess and represent non-photographic media types, such as drawings and animations, for effective NSFW detection.
- Train models using transfer learning and fine-tuning approaches to adapt existing NSFW detection architectures to handle drawings and animations.
- Evaluate the performance of the enhanced models using diverse datasets containing drawings, animations, and photographs to ensure robustness and generalization capability.

Objective 2: Expand NSFW Detection to Assess Degrees of Sexiness

- Define metrics and annotation guidelines to quantify the degree of sexiness in NSFW content, considering factors such as nudity, suggestive poses, and contextual cues.
- Augment existing NSFW detection models with additional layers or modules to assess and classify the level of sexual suggestiveness in detected NSFW content.

- Conduct extensive validation experiments to assess the accuracy and reliability of the sexiness assessment capabilities integrated into the NSFW detection models.

Objective 3: Enable Responsible Deployment of Enhanced NSFW Detection Models

- Investigate ethical considerations surrounding NSFW content detection and sexiness assessment, addressing issues such as privacy, bias, and user consent.
- Develop guidelines and best practices for the responsible deployment of NSFW detection models in online platforms, emphasizing transparency, fairness, and user empowerment.
- Collaborate with industry stakeholders and regulatory bodies to promote the adoption of ethical standards and guidelines for NSFW content moderation in AI-driven systems.

Methodology:

1. Data Collection:

- Gather diverse datasets containing NSFW images, including photographs, drawings, animations, and images with varying degrees of sexual content.
- Ensure proper annotation of the datasets to categorize images based on their NSFW content type (e.g., nudity, suggestive poses) and degrees of sexiness.

2. Data Preprocessing:

- Normalize and standardize the collected datasets to ensure consistency in image quality, resolution, and format.
- Implement data augmentation techniques to increase the diversity of the training data and improve model robustness.

3. Model Selection:

- Choose a suitable base model architecture for NSFW detection, considering factors such as model complexity, computational efficiency, and performance on similar tasks.
- Evaluate candidate models on benchmark datasets to assess their baseline performance and suitability for the task.

4. Transfer Learning and Fine-tuning:

- Initialize the selected model with pre-trained weights from a relevant source (e.g., NSFWJS) to leverage existing knowledge.
- Fine-tune the model on the collected NSFW image datasets, focusing on adapting the model to detect NSFW content in drawings, animations, and assess degrees of sexiness.

5. Sexiness Assessment Module Integration:

- Design and implement a sexiness assessment module to augment the NSFW detection model, enabling it to classify the level of sexual suggestiveness in detected NSFW content.
- Train the sexiness assessment module using annotated data specifically curated for this task, considering various factors contributing to the perception of sexiness.

6. Model Training and Validation:

- Train the integrated NSFW detection and sexiness assessment model using the augmented dataset, employing techniques such as cross-validation and early stopping to prevent overfitting.
- Validate the trained model on separate validation datasets to assess its performance in detecting NSFW content in drawings, animations, and accurately assessing degrees of sexiness.

7. Performance Evaluation:

- Evaluate the performance of the trained model using appropriate evaluation metrics, including precision, recall, F1-score, and area under the receiver operating characteristic curve (AUC-ROC).

- Conduct comparative analysis with baseline models and existing state-of-the-art NSFW detection systems to benchmark the performance improvements achieved.

8. Ethical Considerations:

- Address ethical considerations related to NSFW content detection and sexiness assessment, including privacy concerns, potential biases, and user consent.
- Implement measures to mitigate bias and ensure fairness in model predictions, such as bias correction techniques and diversity-aware training strategies.

9. Deployment and Documentation:

- Deploy the trained NSFW detection model and sexiness assessment module in a controlled environment for further testing and validation.
- Document the entire methodology, including data collection procedures, model architecture, training protocols, evaluation results, and ethical considerations, for transparency and reproducibility.

Background

Web archiving (WA) is a vital field focused on preserving segments of the World Wide Web (WWW) to ensure the retention of valuable information for researchers, historians, and the public. Utilizing web crawlers like Heritrix [1], web archives automate the collection process, capturing diverse content types such as HTML pages, cascading style sheets, JavaScript files, images, videos, and associated metadata. However, selecting which content to preserve presents a challenge due to storage limitations and the exponential growth of web data.

Numerous WA initiatives tackle this challenge, each with its scope and focus. Some archives, like the European Commission Historical Archives [1] and the UK Web Archive [2], concentrate on preserving specific types of web pages or entire national top-level domains. Others, such as the Internet Archive [3], aim to preserve the entire web.

NSFW JS WA initiative, is dedicated to preserving the top-level domain and associated web pages. Beyond preservation, it serves as a research infrastructure, offering open access to its searchable contents, ensuring the accessibility of historically significant information.

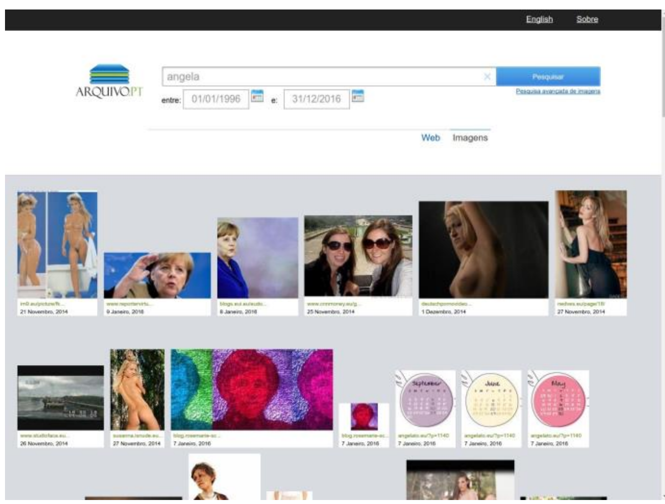
However, the utility of web archives hinges on users' ability to retrieve desired information effectively. To enhance retrieval capabilities, Arquivo.pt provides a full-text search system for its data. Efforts to advance Web Archiving Information Retrieval (WAIR) have explored leveraging temporal information [2] to improve search functionality.

Following this trajectory, Arquivo.pt is developing an Image Search Service (ISS) to bolster its search capabilities further. This service empowers users to retrieve images from NSFW JS contents using natural language queries, thereby enhancing access to visual information within the archive.

A. The Challenge

The internet hosts an immense volume of visual content, primarily stored as graphic files. Within this vast repository lies a subset of content deemed Not Suitable For Work (NSFW) for the majority of users due to its offensive or explicit nature, including images containing nudity, violence, and pornography. Such exposure is particularly concerning for children and young individuals. The NSFW JS Image Search Service (ISS) retrieves images based on queries that match against the filename, alternative text, and surrounding text of an image as presented on a webpage. However, given the diverse and dynamic nature of the web, there are no guarantees that seemingly innocuous queries will not yield results containing offensive content. For example, a website compromised by web spam could unexpectedly return offensive content. This issue is illustrated in Figure 1, where a seemingly harmless query term "angela" returned results with potentially offensive content.

The detection and filtering of NSFW content, specifically nudity and pornography, from the resources archived by Arquivo.pt, necessitates a binary classification task. However, this task is rife with challenges stemming from the sheer scale (billions of images) and diversity (ranging from small to very large images, computer-generated graphics to natural images, etc.) of the image contents within the archived web pages.



Background and Related Work

The task of automatically identifying NSFW (Not Safe For Work) content from images and multimedia has been extensively studied in the literature [3], [4], [5], [6], with numerous research efforts aimed at providing and enhancing methods for content classification. This section provides an overview of the techniques employed for NSFW content classification.

A. Skin Detection Methods

Early methods [6] addressing this problem relied on skin-detection algorithms to identify regions of interest in images, subsequently analyzing the features of these skin regions to determine their pornographic nature. One example of such methods is the POESIA filter [7], an open-source implementation of a skin-color-based filter. However, the performance of these methods is contingent upon the accuracy of the skin detection algorithm and the features extracted, often resulting in high false positive rates, particularly in images featuring beach or sports activities.

B. Bag of Visual Words Methods

Another approach yielding promising results in image classification involves the use of Bag of Visual Words (BoVW) techniques [3]. These methods extract visual features from images, represented as "words" akin to the Bag-of-Words (BOW) approach in text document classification [8]. A vocabulary vector is constructed based on the occurrence of these visual words, typically derived from detecting keypoints or local descriptors such as Scale-Invariant Feature Transform (SIFT) variations. Subsequently, a classifier trained on these representations is used to classify image content as pornographic or non-pornographic.

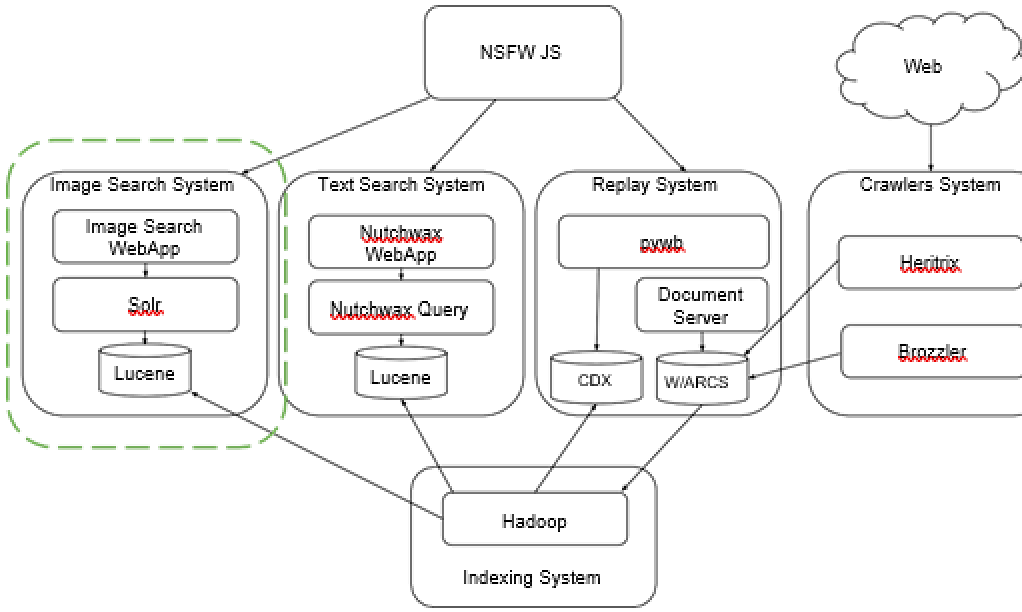
C. Neural Network Based Methods

Neural network-based methods have emerged as powerful tools for image classification tasks. Artificial neural networks (ANNs) are computing systems inspired by biological neural networks, composed of simple structures mapping input values to output values. Networks with numerous neurons and layers are termed Deep Neural Networks (DNNs).

For image classification, input data to the network consists of pixel values from the image. Depending on the architecture and objective, each neural network requires fixed-size input data. A common input data size for color image classification is a 3x256x256 array, where each array unit corresponds to a pixel value from a specific RGB channel. Each array value serves as a feature, contributing to a feature vector with a dimensionality of 196,608. Neural networks, particularly Deep Neural Networks (DNN), represent the forefront of research in machine learning. These networks are adept at learning representations from data, emphasizing the acquisition of successive layers of increasingly meaningful features. Characterized by their depth, parameters, layers, and neurons, DNNs have demonstrated remarkable success, particularly in image recognition tasks. Convolutional Neural Networks (CNNs), a specific type of DNN, have garnered significant attention for their efficacy in image recognition.

CNNs have exhibited state-of-the-art performance in various image recognition tasks, including NSFW (Not Safe For Work) image classification. Several CNN architectures have been developed, each achieving improved accuracy on benchmarks like the ImageNet classification challenge. Notable architectures include ALexNet, ZF Net, GoogLeNet, and Residual Networks.

DL (Deep Learning) techniques seamlessly combine feature extraction and classification, reducing the need for manual feature selection. However, these approaches necessitate substantial amounts of training data and computational infrastructure. Fortunately, datasets like ImageNet and pre-trained models are readily available, simplifying the implementation of DL models. For instance, Yahoo!'s OpenNSFW model offers an open-source solution for identifying NSFW images, specifically pornography, providing developers with pre-trained models for classification purposes. Despite their efficacy, concerns remain regarding the availability of training data and the infrastructure requirements associated with DL techniques.



Proposed Approach

This section outlines the proposed solution for classifying NSFW images. Subsection IV-A provides an overview of the CNN models utilized for this task, while Subsection IV-B delves into the integration of the proposed solution into the Arquivo.pt infrastructure.

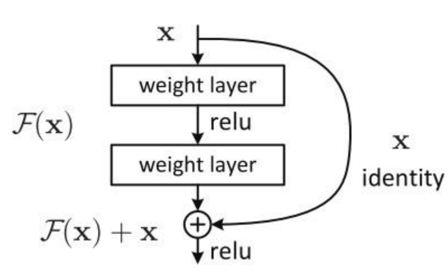
% Amount	Number of URL	mime-types
79.60	237 966 251	text/html
10.03	29 997 689	image/jpeg
2.45	7 328 179	image/png
0.77	2 305 834	application/pdf
0.76	2 271 770	application/javascript
0.72	2 145 481	text/xml
0.62	1 869 902	application/rss+xml
0.62	1 861 165	application/json
0.60	1 820 236	image/gif
0.58	1 783 630	text/css
3.25	1 783 630	all others

% Amount	Number of URL	mime-types
75.59	30 145 538	image/jpeg
18.37	7 328 179	image/png
4.56	1 820 236	image/gif
0.98	389 839	image/svg+xml
0.34	135 720	image/x-icon
0.09	38 577	image/pjpeg
0.06	22 717	image/bmp

A. Deep Neural Network Models

In this study, we conducted experiments using two distinct architectures of Deep Neural Networks (DNN): ResNet [17] and SqueezeNet networks [24]. These networks delineate a hypothesis space, which confines the realm of possibilities for solving a specific problem. It is within this constrained space that we endeavor to uncover a useful representation of the input data, subsequently mapped to the desired output.

DNNs pose challenges in training due to issues such as the vanishing and exploding gradient problem [25]. This problem arises when the gradient values propagated back to the initial layers undergo repeated multiplications, potentially resulting in infinitesimally small or large gradient values. Consequently, network learning may stagnate, particularly in the early layers. Various strategies have been devised to tackle these issues, including normalized initialization [25], rectifiers [26], and normalization layers [27]. With the increasing depth of DNNs, certain network nodes may become saturated, leading to accuracy degradation (not attributed to overfitting). ResNet addresses this degradation problem by introducing shortcut connections between layers, as illustrated in Figure 3.



The fundamental concept behind ResNet is that deeper networks should not degrade network training performance [17]. Stacking identity mappings atop the existing network should yield comparable performance. The hypothesis is that enabling the stacked layers to fit a residual mapping is easier than directly fitting the desired underlying mapping. Research has demonstrated that ResNet networks are easier to optimize and achieve higher accuracy due to their increased depth.

SqueezeNet, on the other hand, was designed to have a smaller architecture while preserving the accuracy of larger models [24]. This network boasts fewer parameters and layers, offering advantages such as more efficient distributed training, reduced overhead in exporting new models, and compatibility with Field-Programmable Gate Arrays (FPGAs) and embedded circuits. The authors achieve these objectives through various strategies, including replacing 3x3 filters with 1x1 filters, decreasing the number of input channels to 3x3 filters, and downsampling late in the network to ensure convolution layers have large activation maps.

The architecture of SqueezeNet incorporates Fire Modules [24], illustrated in Figure 4. These modules consist of a squeeze convolution layer with 1x1 filters to compress incoming data, as well as an expand convolution layer with a mix of 1x1 and 3x3 convolution filters to augment the depth of the data.

In our solution, two critical parameters of the DNN need consideration: the loss function and the optimizer employed for network training. The choice of a suitable loss function, also known as the objective function, is pivotal as it provides feedback on the model's performance. It guides the network in minimizing the loss effectively, thereby enhancing its predictive capabilities.

The loss function, or log loss, evaluates the performance of a classification model by measuring the discrepancy between predicted probabilities (ranging from 0 to 1) and actual labels. Cross-entropy loss increases as the predicted probability deviates from the true label. Ideally, a perfect model would yield a log loss of 0, indicating precise predictions. Thus, minimizing cross-entropy equates to maximizing the log likelihood of our data, serving as a direct measure of the predictive capacity of our model.

The optimizer governs the adjustment of network weights based on predictions and the loss function. Through incremental updates, the optimizer progressively reduces the loss score, refining the accuracy of network predictions. Optimizers typically fall into two categories: first-order and second-order algorithms. First-order optimization algorithms, such as Gradient Descent (GD), utilize the gradient of the loss function to minimize network loss. Conversely, second-order algorithms employ the second-order derivative to minimize loss, enabling assessment of the function's curvature. While second-order algorithms provide valuable insights, they entail higher computational costs, making first-order optimizers more commonly used. In this study, we employ a Stochastic Gradient Descent (SGD) algorithm [29].

I. Experimental Evaluation

This section presents the experimental evaluation conducted on the proposed solution, organized as follows. Section V-A describes the construction of the initial evaluation dataset. Section V-B outlines the software and hardware platforms utilized for the evaluations, along with standard evaluation metrics. Image preprocessing steps are detailed in Section V-C, followed by the presentation of experimental results in Section V-D. The integration of the solution is reported in Section V-E, and Section V-F addresses the experimental results of model improvements.

A. Constructing the Initial Evaluation Dataset

Due to the limited availability of datasets for NSFW classification, particularly those containing only raw content locations, we opted to create our own dataset. This dataset comprises 17,655 images manually labeled from Arquivo.pt, with 8,273 labeled as NSFW and 9,382 as SFW (see Table III). An additional

18,626 images remain unlabeled. Notably, the number of NSFW images in Arquivo.pt is considerably smaller compared to SFW images, posing a challenge in dataset compilation to ensure balanced classes.

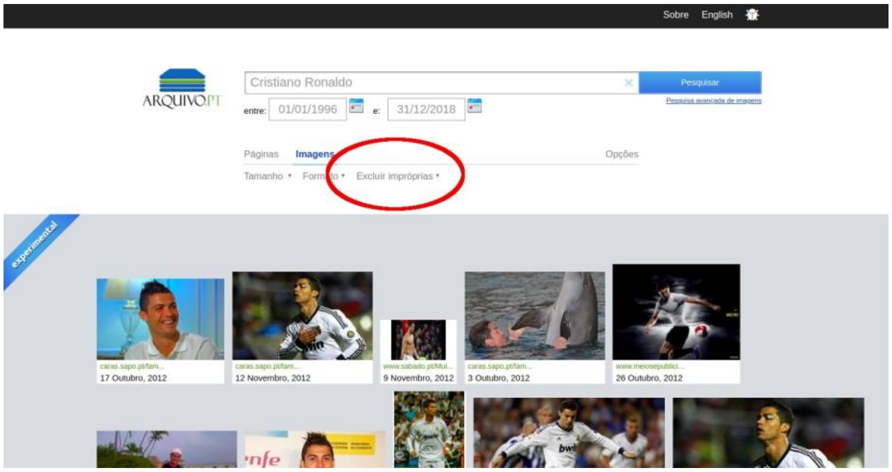
To gather NSFW images, two main methods were employed. The first method involved utilizing a crowdsourcing platform where users labeled images returned from queries processed through Arquivo.pt's search functionality. The second method utilized Arquivo.pt Text Search API to retrieve web pages containing NSFW-related terms such as "porn," "blowjob," or "fuck," which are typically associated with NSFW content [31]. Images extracted from these web pages constituted a significant portion of the dataset, totaling 18,000 images.

B. Evaluation Setup and Metrics

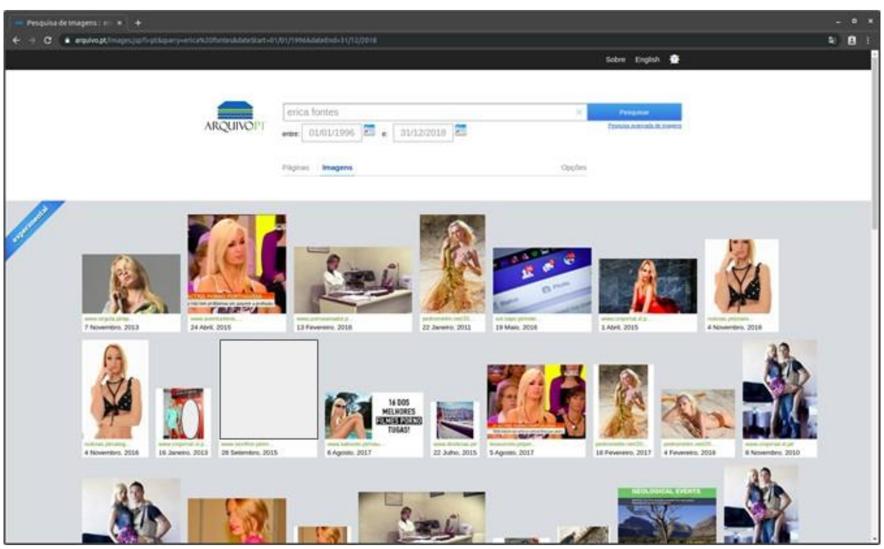
The models were evaluated using standard hardware configurations, including a common laptop with 8 GB RAM, a GeForce GTX 860M GPU, and an Intel(R) Core(TM) i7-4710HQ CPU @ 2.50GHz. Additionally, evaluations were conducted on server-class hardware available at Arquivo.pt infrastructure, featuring a Dell PowerEdge R730xd model with 256 GB RAM and an Intel(R) Xeon(R) CPU E5-2620 v3 @ 2.4 GHz.

To compare the models, various evaluation metrics were employed, including the Area Under the Curve (AUC) computed over the Receiver Operating Characteristic (ROC) curves. This metric, commonly used for binary classifiers, evaluates classifier performance across varying threshold values, facilitating assessment before selecting a specific threshold. Additionally, metrics such as accuracy, precision, recall, and F-score were calculated, providing comprehensive insight into model performance.

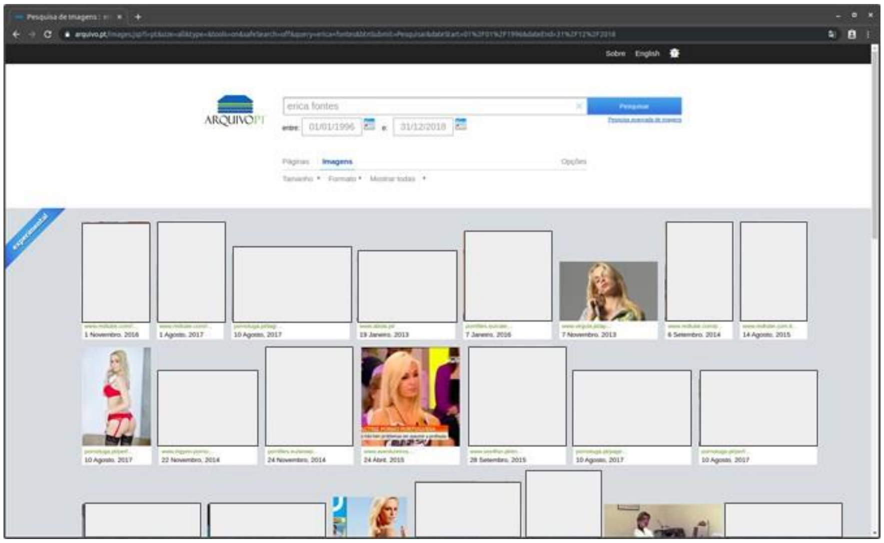
NSFW JS hosts over six million preserved resources, offering users access to a vast repository of information. To facilitate information retrieval among these contents, Arquivo.pt provides text-based search tools and is currently developing an Image Search System (ISS) to enhance search capabilities. However, given the abundance of visual content on the web and within Arquivo.pt, there's a risk of encountering offensive material such as nudity, violence, or pornography.



Integration of Image Classification Interface with Query Term 'Cristiano Ronaldo'. The user selected the NSFW filter option, as indicated by the red ellipse.



Integration of Image Filtering Interface with Query Term 'Erica Fontes', utilizing the NSFW classifier. The gray rectangular box, denoted by the third image in the second row, highlights NSFW content that was misclassified.



Integration of Image Filtering Interface with Query Term 'Erica Fontes', without utilizing the NSFW classifier. The gray rectangular boxes indicate NSFW content that has been both highlighted and hidden.

Conclusion

To address this issue, a solution was developed and integrated into Arquivo.pt ISS to automatically identify images containing such content. This solution utilizes a Convolutional Neural Network (CNN) to classify images and provides the classification result to the ISS, enabling the system to filter out NSFW (Not Suitable For Work) content from search results.

Two pre-trained deep neural networks were evaluated using a dataset comprised of Arquivo.pt images. Subsequently, a fine-tuning process was conducted to enhance the accuracy of the models in identifying NSFW content. The initial evaluation yielded 93% accuracy for the ResNet model and 72% for the SqueezeNet model. Following fine-tuning, the accuracy improved to 94% for ResNet and 89% for SqueezeNet.

The optimized model was integrated into the ISS indexing workflow, utilizing a message queue to distribute and scale the classification workload across multiple workers. The proposed solution is currently operational on https://nsfwjs.com/

Future work includes improving model accuracy by expanding the dataset, utilizing more recent deep neural network models, and increasing the number of training epochs. Additionally, the classification system could be extended to include multiple NSFW categories, such as strict, moderate, or off, similar to the categorization offered by search engines like DuckDuckGo. Furthermore, the system can be expanded to identify other types of NSFW content, such as violence and offensive symbols.

References List:

1. Gordon Mohr, Michael Stack, Igor Ranitovic, Dan Avery, and Michele Kimpton. Introduction to Heritrix, an archival quality web crawler. In 4th International Web Archiving Workshop (IWAW04), Bath, UK, 2004.

2. Miguel Costa. Information Search in Web Archives. PhD thesis, Faculty of Sciences of the University of Lisbon, December 2014.

3. T. Deselaers, L. Pimenidis, and H. Ney. Bag-of-visual-words models for adult image classification and filtering. In 2008 19th International Conference on Pattern Recognition, pages 1–4, Dec 2008.

4. D.A. Forsyth and M.M. Fleck. Automatic detection of human nudes. International Journal of Computer Vision, 32(1):63–77, Aug 1999.

5. T. Lindeberg. Scale Invariant Feature Transform. Scholarpedia, 7(5):10491, 2012. revision #153939.

6. Huicheng Zheng, Mohamed Daoudi, and Bruno Jedynak. Blocking Adult Images Based on Statistical Skin Detection. Electronic Letters on Computer Vision and Image Analysis, 4(2):1–14, 2004.

7. Daoudi Mohamed. POESIA - Filtering Software @ONLINE, January 2018.

8. C. Manning, P. Raghavan, and H. Schu¨tze. Introduction to information retrieval. Cambridge University Press, 2008.

9. Gu Jiuxiang et al. Recent advances in convolutional neural networks. CoRR, abs/1512.07108, 2015.

10. Alex Krizhevsky, IIya Sulskever, and Geoffret E Hinton. ImageNet Classification with Deep Convolutional Neural Networks. Advances in Neural Information and Processing Systems (NIPS), pages 1–9, 2012.

11. Y. Sun, B. Xue, M. Zhang, and G. G. Yen. Evolving deep convolutional neural networks for image classification. IEEE Transactions on Evolutionary Computation, 24(2):394–407, 2020.

12. Dmirty Zhelonkin and Nikolay Karpov. Training effective model for real-time detection of nsfw photos and drawings. In International Conference on Analysis of Images, Social Networks and Texts, pages 301–312. Springer, 2019.

13. Ruolin Zhu, Xiaoyu Wu, Beibei Zhu, and Liuyihan Song. Application of pornographic images recognition based on depth learning. In Proceedings of the 2018 International Conference on Information Science and System, pages 152–155, 2018.

14. Stanford University. ImageNet, http://www.image-net.org/, January 2018.

15. Matthew D. Zeiler and Rob Fergus. Visualizing and understanding convolutional networks. In David Fleet, Tomas Pajdla, Bernt Schiele, and Tinne Tuytelaars, editors, Computer Vision – ECCV 2014, pages 818–833, Cham, 2014. Springer International Publishing.

16. Christian Szegedy et al. Going deeper with convolutions. CoRR, abs/1409.4842, 2014.

17. Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. Deep Residual Learning for Image Recognition. In IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pages 770–778, 2016.

18. R. Duda, P. Hart, and D. Stork. Pattern classification. John Wiley & Sons, 2nd edition, 2001.

19. F. Escolano, P. Suau, and B. Bonev. Information theory in computer vision and pattern recognition. Springer, 2009.

20. Jay Mahadeokar. Open NSFW model code @ONLINE, January 2018.