Mr. R.A.Ghadage, 2024, 12:5
ISSN (Online): 2348-4098
ISSN (Print): 2395-4752

International Journal of Science, Engineering and Technology

An Open Access Journal

Deepfake Detection in Medical Images

Professor Mr. R.A.Ghadage, Vrushali Anil Zinj, Priya Prashant Sarode

Department of Computer Engineering, Ahmednagar Jilha Maratha Vidya Prasarak Samaj's Shri Chhatrapati Shivaji Maharaj College of Engineering, Nepti, Ahmednagar

Abstract- The challenge of deepfake detection in medical imaging by leveraging the Mask R-CNN algorithm. Deepfakes, generated using advanced AI, manipulate images and videos, posing significant threats across industries, including healthcare. Altered medical images can lead to misdiagnoses, treatment delays, or inappropriate interventions, putting patients at risk. The ability to identify such manipulations is critical for maintaining trust in medical diagnoses. Hospitals relying on compromised data may experience disruptions, financial losses, and legal complications. This project aims to develop an efficient deep learning-based system to detect these synthetic alterations. By using the Mask R-CNN framework, the proposed solution seeks to accurately locate and flag tampered regions within medical images. The model enhances patient safety by ensuring the reliability of diagnostic data. Ultimately, this approach offers a safeguard for healthcare institutions against the dangers posed by deepfake technology.

Keywords- Deepfake Detection, Medical Imaging, Mask RCNN Algorithm, Artificial Intelligence, Deep Learning, Image Manipulation

I. INTRODUCTION

The project titled "Deepfake Detection in Medical Imaging Using Mask R-CNN Algorithm" focuses on identifying manipulated medical images, specifically addressing the challenges associated with deepfake technology. Deepfake images—produced by advanced machine learning models—can add, remove, or alter medical features such as tumors, leading to serious diagnostic errors.

Manipulated medical data can mislead healthcare professionals, resulting in inaccurate diagnoses and treatment plans, with potentially life-threatening consequences. The goal of this project is to develop a reliable system that distinguishes between real and altered medical images, ensuring the authenticity of diagnostic information and protecting healthcare institutions from the risks posed by such manipulations. The core technology behind this project is the Mask R-CNN algorithm, known for its ability to detect and segment objects within images at a high precision. In this context, the model is trained to detect two types of deepfake manipulations: tumors falsely added to images and existing tumors removed from them. By identifying these changes accurately, the system ensures that healthcare professionals can rely on the integrity of medical reports, preventing costly errors and safeguarding patient outcomes.

II. PROBLEM STATEMENT

In recent years, deepfake technology has advanced rapidly, enabling the creation of highly convincing manipulated media, including medical images. In the healthcare sector, the use of such synthetic modifications presents a severe threat to diagnostic integrity. Malicious actors can use deepfake techniques to either add non-existent tumors or remove real tumors from medical scans, potentially leading to misdiagnoses, incorrect treatment plans, and compromised patient care.

III. PROPOSED SYSTEM

The proposed system leverages the Mask R-CNN algorithm to detect deepfake alterations in medical images, ensuring the authenticity of diagnostic data. Mask R-CNN, known for its precise object detection and segmentation capabilities, is employed to identify manipulations such as added or removed tumors. The system will process medical images, including MRI and CT scans, and classify them into four categories: real tumor detected, no tumor detected, tumor added (deepfake), and tumor removed (deepfake). Once an image is uploaded, the algorithm analyzes it by localizing regions of interest, extracting image features, and comparing them with genuine patterns based on pre-trained models. The system will offer a user-friendly interface where healthcare professionals can upload images and receive real-time feedback on their authenticity. It also provides an optional training module for continuously updating the model with new datasets to enhance detection accuracy over time. The objective is to safeguard medical institutions from synthetic media threats, ensuring that patient treatment decisions are based on accurate data. By preventing deepfake manipulations, the proposed system will help maintain diagnostic integrity, improve patient outcomes, and minimize the risks associated with fraudulent medical images.

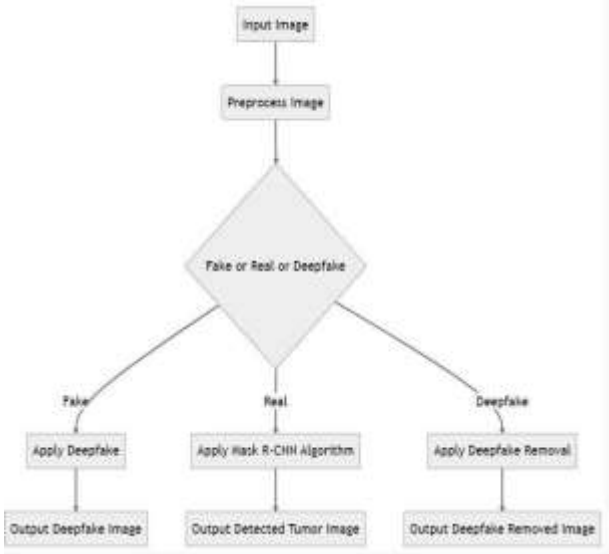

Fig 1: System Architecture

IV. SYSTEM ARCHITECTURE

Module 1: Image Acquisition and Preprocessing
This module handles image input by allowing users (radiologists or healthcare professionals) to upload medical scans such as MRIs or CT images. The preprocessing step involves resizing, normalization, and noise reduction to ensure the image is in the proper format and quality for analysis. Data augmentation techniques may also be applied to improve the robustness of the model during the training phase.

Module 2: Deepfake Detection Module using Mask R-CNN
This module is the core of the system, utilizing the Mask RCNN algorithm to analyze the uploaded images. It performs object detection, segmentation, and classification to determine if a tumor is real or manipulated. The model identifies regions where a tumor is added (deepfake) or removed (deepfake) and distinguishes these from authentic medical images. Results are categorized into: Brain Tumor Detected (Real) No Tumor Detected (Real) Tumor Added (Deepfake) Tumor Removed (Deepfake).

Module 3: User Interface Module
This module provides a simple and interactive interface where users can upload medical images, view results, and interact with the detection system. The UI displays the uploaded image and detection outcome in real-time. It ensures ease of use, making it accessible for medical professionals. It may also include a training section to allow administrators to add new datasets and enhance the model's accuracy.

Module 4: Training and Model Update Module
This module enables the continuous learning and improvement of the system by incorporating new data. Healthcare professionals or developers can train the model with additional medical images to enhance its performance. This module manages dataset preparation, model retraining, and validation, ensuring that the system adapts to new challenges in deepfake detection. It ensures that the system remains up-to-date with evolving

deepfake techniques, making it more resilient over time.

V. WORKING OF SYSTEM

The system operates by leveraging the Mask R-CNN algorithm to detect and classify deepfake manipulations in medical imaging. Upon uploading an image, such as an MRI or CT scan, the model initiates a multi-step analysis. First, the image is preprocessed to enhance clarity and prepare it for detailed examination. The Mask R-CNN then segments the image into distinct regions, enabling it to identify both the presence of tumors and any alterations made to them. Specifically, the algorithm focuses on detecting two manipulation types: the addition of false tumors and the removal of existing ones. Each segment is evaluated against trained data to ascertain its authenticity. The results are categorized into four outputs: "Brain Tumor Detected (Real)," indicating a genuine tumor; "No Tumor Detected," confirming the image's integrity; "Tumor Added (Deepfake)," signaling the presence of a fabricated tumor; and "Tumor Removed (Deepfake)," which indicates that an existing tumor has been manipulated. This systematic approach ensures a reliable assessment of medical images, facilitating informed decision-making by healthcare professionals and safeguarding patient safety.

VI. SYSTEM DESIGN

The system design for the "Deepfake Detection in Medical Imaging Using Mask R-CNN Algorithm" project consists of several integrated components working together to ensure accurate detection of manipulated medical images. At the forefront is a user-friendly interface that allows healthcare professionals to easily upload medical images, such as MRIs or CT scans, for analysis. Once an image is uploaded, it undergoes preprocessing, including normalization, resizing, and noise reduction, to enhance its quality. The core of the system is the Mask R-CNN model, which utilizes a Region Proposal Network (RPN) to identify potential areas of interest within the image, followed by a segmentation branch that distinguishes between real tumors, added tumors, and removed tumors. After analysis, the model categorizes the results into four distinct outputs: "Brain Tumor Detected (Real)," "No Tumor Detected," "Tumor Added (Deepfake)," and "Tumor Removed (Deepfake)," presenting this information clearly on the interface.

Additionally, an optional training module allows users to input new datasets, enabling continuous improvement of the model's accuracy. To address data security and privacy concerns, the system incorporates robust measures to protect sensitive medical information, ensuring compliance with data protection regulations. Overall, this comprehensive system design prioritizes usability, accuracy, and security, providing a reliable solution for detecting deepfake manipulations in medical imaging.

VII. Technical Requirements

1. Hardware Requirements

Processing Unit

A powerful GPU (Graphics Processing Unit) is essential for training and running the Mask R-CNN model efficiently. Recommended options include NVIDIA RTX 2080 Ti or better, which support CUDA for accelerated computing.

RAM

A minimum of 16 GB of RAM is required to handle large datasets and enable smooth multitasking during the image processing and training phases.

Storage

At least 1 TB of SSD (Solid State Drive) storage is necessary to accommodate the operating system, software tools, datasets, and model outputs, providing fast read/write speeds for data access.

2. Software Requirements

Operating System: The system should run on a compatible OS such as Windows 10, Ubuntu 22.04, or any other Linux distribution that supports deep learning frameworks.

Development Environment: Python 3.11 or higher is required for developing the application, along with essential libraries such as TensorFlow or

PyTorch for implementing the Mask R-CNN algorithm.

Image Processing Libraries: OpenCV and scikit-image libraries are recommended for image preprocessing tasks, including normalization and noise reduction.

Web Framework: Tkinter should be utilized for developing the web application, enabling user interface functionality and server-side processing.

Future Scope

The future scope of the "Deepfake Detection in Medical Imaging Using Mask R-CNN Algorithm" project presents numerous opportunities for advancement and integration within the healthcare sector. One key area for development is the expansion of manipulation types, enabling the detection of alterations beyond tumors, such as changes in anatomical structures in various imaging modalities like X-rays and ultrasounds. Additionally, integrating the detection system with advanced imaging technologies could facilitate real-time analysis during diagnostic procedures, improving workflow efficiency for healthcare professionals. Collaborating with hospitals and clinics to implement the system in clinical settings would provide valuable real-world data for continuous refinement and validation, while enhancing the Mask R-CNN algorithm through techniques like transfer learning could improve detection accuracy and speed. Comprehensive training programs for healthcare professionals would ensure effective interpretation of results and implications of detected manipulations. Transitioning to a cloud-based platform would allow for greater scalability, enabling broader access without extensive local infrastructure and promoting collaborative research efforts. Addressing regulatory compliance and ethical considerations will also be crucial for the system's acceptance in clinical practice. Lastly, educating patients about the risks of manipulated medical images and the importance of detection technologies could foster trust in the healthcare system and encourage the adoption of such innovations. Overall, these developments can transform the project into a comprehensive solution that enhances the integrity and safety of diagnostic practices in healthcare.

VIII. CONCLUSION

In conclusion, the "Deepfake Detection in Medical Imaging Using Mask R-CNN Algorithm" project addresses a critical challenge in healthcare by providing a reliable solution for identifying manipulated medical images. The rise of deepfake technology poses significant risks, potentially leading to misdiagnoses and compromised patient safety. By leveraging the advanced capabilities of the Mask R-CNN algorithm, the system can accurately detect alterations in medical images, such as the addition or removal of tumors, ensuring the integrity of diagnostic information. The design and implementation of the system prioritize usability, allowing healthcare professionals to seamlessly upload and analyze medical images while receiving clear, actionable results. The inclusion of a training module facilitates continuous improvement of the model, adapting to new data and enhancing its accuracy over time. Moreover, robust security measures ensure the protection of sensitive medical information, adhering to privacy regulations.

Acknowledgement

We would like to take this opportunity to thank all the people who were part of this seminar in numerous ways, people who gave un-ending support right from the initial stage. In particular, we wish to thank Prof. R. A. Ghadhge as an internal project guide who gave their co-operation timely and precious guidance without which this project would not have been a success. We thank them for reviewing the entire project with painstaking efforts and more of her, unbanning ability to spot the mistakes.

REFERENCES

[1]. Ahmed, M., Islam, M., Das, M., & Khan, A. (2021). Mapping and situation analysis of basic WASH. PLOS ONE, 11(16), 1-12.

[2]. Akter, T., & Ali, A. (2014). Factors influencing knowledge and practice of hygiene in Water,

Sanitation and Hygiene (WASH) programme areas of Bangladesh Rural Advancement Committee. Rural and Remote Health, 14(2628), 1-11.

[3]. Badhan, M. A., Roy, B., & Sifat, S. A. (2017). Water supply and sanitation situation of Kalyanpur slum area in Dhaka. International Journal of Natural and Social Sciences, 2(4), 54-59.

[4]. Bangladesh Bureau of Statistics (BBS) and UNICEF Bangladesh. (2014). Multiple indicator cluster survey. Dhaka: Bangladesh Bureau of Statistics (BBS) and UNICEF Bangladesh.

[5]. Gafan, N., Kpozèhouen, A., & Dégbey, C. (2022). Household access to basic drinking water sanitation and hygiene facilities: secondary analysis of data from the demographic and health survey V, 2017–2018 . BMC Public Health, 22(1345), 1-17.

[6]. Goyanka, R. (2021). Burden of water, sanitation and hygiene related diseases in India:. Clinical Epidemiology and Global Health, 1-8.

[7]. Khan, M. (2022). Livelihood, WASH related hardships and needs assessment of climate migrants: evidence from urban slums in Bangladesh. Heliyon(8), 1-10.

[8]. Meshi, E. B., Nakamura, K., Seino, K., & Alemi, S. (2022). Equity in water, sanitation, hygiene, and waste management services in. Public Health in Practice(4), 1-9.

[9]. Misha, F., & Munshi, S. (2016). Bangladesh Priorities: Poverty. Copenhagen: Copenhagen Consensus Center.

[10]. Mustafa, S., Jamil, K., Lifu, Z., & Girmay, M. B. (2022). Does Public Awareness Matter to Achieve the UN's Sustainable Development Goal 6: Clean Water for Everyone? Journal of Environmental and Public Health.

[11]. Qurat-ul-Ann, D.-R., & Bibi, M. (2022). Household Multidimensional Water, Sanitation, and Hygiene Poverty in Pakistan. Research Square.

[12]. Rana, M. (2009). Status of water use sanitation and hygienic condition of urban slums: A study on Rupsha Ferighat slum, Khulna. Desalination ., 322-328.

[13]. Rana, S., & Ghosh, H. (2016). Water Supply and Sanitation Status of Haryzon Polly Dwellers at Natunbazar Area in Mymensingh District. Environmental Science and Natural Resources, 143-146.

[14]. Roy, B., & Mohanta, S. (2017). Water supply and sanitation status of low income area in Mymensingh District. International Journal of Natural and Social Sciences, 4(3)(2313-4461), 43-48.

[15]. Safo-Adu, G., Owusu-Adzorah, N., & Essiam, C. (2022). Environmental Sanitation and Hygienic Conditions. Public Health Research, 2(12), 51-59.

[16]. Sarkar, S., & Bharat, G. (2021). Achieving Sustainable Development Goals in water and sanitation sectors in India. Journal of Water, Sanitation and Hygiene for Development, 693–705.

[17]. Shaibur, M. R. (2019). Water Supply and Sanitation Status in Jashore Municipality, Bangladesh. Environmental and Biological Research, 12-21.

[18]. Swain, P., & Pathela, S. (2016). Status of sanitation and hygiene practices in the context of "Swachh Bharat Abhiyan" in two districts of India. International Journal of Community Medicine and Public Health, 3140-3146.

[19]. The Lancet. (2021). The Lancet Commission on water, sanitation and hygiene, and health. The Lancet, 1469-1470.

[20]. United Nations Children's Fund (UNICEF). (2019). Global Framework for Urban Water, Sanitation and Hygiene. New York: United Nations Children's Fund (UNICEF).

[21]. WHO. (2018). WHO Water, Sanitation and Hygiene strategy 2018-2025. Switzerland.

[22]. Wikipedia. (2023, May 27). Wikipedia. Retrieved from Wikipedia: https://en.wikipedia.org/wiki/Mymensingh

[23]. World Health Organization . (2019). Water, sanitation, hygiene and health: a primer for health professionals. Geneva: World Health Organization .

[24]. World Health Organization and UNICEF. (2006). Meeting the MDG drinking water and sanitation target : the urban and rural challenge of the decade. Geneva: World Health Organization.