Journal of Production, Operations and Systems Management

e-ISSN: 1984-2430

Open Acess

Responsible Editor: PhD Hermes Moretti Ribeiro da Silva

Received: 13/08/2023 Approved: 18/12/2023

INVESTIGATING DEEP LEARNING APPLICATIONS IN COMPUTER VISION FOR EFFECTIVE FACIAL MASK DETECTION DURING THE GLOBAL COVID-19 CRISIS

EXPLORANDO O APRENDIZADO PROFUNDO EM VISÃO COMPUTACIONAL PARA EFETIVA DETECÇÃO DE MÁSCARAS FACIAIS NA CRISE GLOBAL DA COVID-19

Wiliam Regone[1],
Luciel Henrique de Oliveira[2]

ABSTRACT

Purpose: This study aims to scrutinize the integration of deep learning algorithms within the sphere of computer vision, with a concentrated focus on proficiently detecting face mask usage amidst the global COVID-19 pandemic.
Theoretical Framework: The research is grounded in the theoretical underpinnings of deep learning, a branch of artificial intelligence, and its application in computer vision. It explores the advancements in machine learning algorithms capable of complex image processing and pattern recognition, essential for identifying face mask usage in various settings.
Methodology/Approach: The research adopts a methodological approach involving the design and development of a deep learning model. This model is trained on a diverse dataset encompassing images of individuals with and without face masks. Python, along with libraries such as OpenCV, Keras, and TensorFlow, forms the backbone of the implementation, facilitating the processing and analysis of image data.
Findings: The study's findings reveal that the developed model demonstrates a high degree of accuracy, with a 99% success rate in test image predictions, showcasing the effectiveness of deep learning in image recognition tasks. This underscores the model's proficiency in identifying face mask usage, a critical factor in controlling the spread of airborne viruses like COVID-19.
Research, Practical & Social Implications: This research contributes significantly to the field of computer vision, offering practical applications in public health monitoring and societal well-being. The model's ability to accurately detect face mask usage paves the way for enhanced pandemic management strategies and reinforces the role of technology in public health initiatives.
Originality/Value: This study innovates within existing research by applying deep learning in computer vision for addressing the COVID-19 crisis. It uniquely focuses on developing technological solutions for efficient and cost-effective monitoring of face mask usage, emphasizing prevention.
Keywords: Computer Vision, Convolutional Neural Network, Face Mask Detection.

RESUMO

Objetivo: Examinar a integração de algoritmos de aprendizado profundo no campo da visão computacional, com foco na detecção eficiente do uso de máscaras faciais durante a pandemia global de COVID-19.
Referencial Teórico: A pesquisa se baseia nos fundamentos teóricos do aprendizado profundo, um ramo da inteligência artificial, e sua aplicação em visão computacional. Explora os avanços em algoritmos de aprendizado de máquina capazes de processar imagens complexas e reconhecimento de padrões, essenciais para identificar o uso de máscaras faciais em diversos contextos.
Metodologia/Abordagem: Foi adotada uma abordagem metodológica envolvendo o desenho e desenvolvimento de um modelo de aprendizado profundo. Este modelo é treinado em um conjunto de dados diversificado que inclui imagens de indivíduos com e sem máscaras faciais. Python, juntamente com bibliotecas como OpenCV, Keras e TensorFlow, formam a espinha dorsal da implementação, facilitando o processamento e análise de dados de imagem.
Resultados: Os achados do estudo revelam que o modelo desenvolvido demonstra um alto grau de precisão, com uma taxa de sucesso de 99% nas previsões de imagens de teste, evidenciando a eficácia do aprendizado profundo em tarefas de reconhecimento de imagens. Isso destaca a competência do modelo em identificar o uso de máscaras faciais, fator crítico no controle da propagação de vírus transmitidos pelo ar como a COVID-19.
Contribuições, implicações práticas e sociais: Esta pesquisa contribui para o campo da visão computacional, oferecendo aplicações práticas no monitoramento da saúde pública e no bem-estar social. A capacidade do modelo de detectar com precisão o uso de máscaras faciais abre caminho para estratégias aprimoradas de gestão de pandemias e reforça o papel da tecnologia em iniciativas de saúde pública.
Originalidade/Valor: Esta pesquisa explora a implementação de técnicas avançadas de aprendizado profundo no campo da visão computacional, visando a mitigação de crises de saúde pública global, como a pandemia da COVID-19. A originalidade é evidenciada na sua abordagem inovadora para atender às demandas críticas impostas por pandemias, através do emprego de soluções tecnológicas. O estudo destaca-se ao desenvolver e propor métodos que são não apenas preventivos e eficientes, mas também economicamente sustentáveis, para o monitoramento do uso de máscaras faciais, uma necessidade vital no contexto de saúde pública atual.
Palavras-chave: Visão computacional; Rede Neural Convolucional; Máscara facial.

DOI:10.15675/gepros.2988

[1] Pontifícia Universidade Católica de Minas Gerais - PUC Minas. Depto de Engenharia de Produção. Av. Pe. Cletus Francis Cox, 1.661, Poços De Caldas, MG, Brasil, e-mail: wiliamregone@gmail.com, Orcid: https://orcid.org/0000-0003-4731-1642

[2] E-mail: luciel@uol.com.br, Orcid: https://orcid.org/0000-0002-0057-1423

1. INTRODUCTION

The COVID-19 pandemic has significantly impacted global economies, environmental dynamics, and social norms, affecting millions worldwide. This unprecedented situation has necessitated the implementation of various public safety measures, among which mask-wearing and social distancing, as recommended by the World Health Organization, are paramount. The manual enforcement of these measures, particularly mask-wearing, poses a significant challenge. In response, this study explores the application of computer vision, augmented by Artificial Intelligence (AI), to assist in monitoring adherence to these safety protocols in public spaces.

This research is guided by the question: How can AI and computer vision effectively monitor the usage of face masks in public areas, thereby aiding in the control and prevention of COVID-19? To address this, the study proposes an AI-based methodology to statistically analyze the compliance with mask-wearing regulations.

Computer vision, an interdisciplinary field, involves the processing and interpretation of digital images to enable high-level understanding by computers. This field has seen substantial growth, primarily fueled by advancements in deep learning, a subset of AI focusing on computational models with multi-layered processing for abstract data representation (Yadav, 2020). Central to this approach are convolutional neural networks (CNNs), inspired by biological neural networks, which process inputs through interconnected artificial neurons and mathematical functions to identify patterns and outputs.

Deep learning facilitates the learning of data representations at various abstraction levels. These models are adept at tasks such as face detection and recognition, object identification, human action classification in videos, and motion tracking. The deep learning approach, characterized by its multi-layered processing, is aptly named for its depth of computational analysis (Apostolopoulos & Mpesiana, 2020).

Given the escalating use of face masks due to the pandemic, with a shift from pollution protection to mandatory public health measures, this research introduces a novel detection system to identify individuals with or without face masks. This system leverages the principles of computer vision and deep learning through CNNs, offering a refined approach to enhance existing public safety protocols.

2. LITERATURE REVIEW

The discipline of computer vision has undergone significant evolution, largely driven by the development and application of neural networks. Central to this advancement are Artificial Neural Networks (ANNs), comprising an intricate web of interconnected neurons designed to optimize processing efficacy. A notable innovation in this field is the Convolutional Neural Network (CNN), distinguished by its neurons' capacity for autonomous optimization through learning. These neurons perform vectorial mathematical operations under non-linear functions, underpinning a multitude of ANNs. CNNs, specifically adept at processing image data, organize their neuronal layers in three dimensions: height, width, and depth, with depth representing the number of classes in a dataset. Contrasting traditional networks, CNN neurons within a layer connect only to a limited region of the preceding layer, as elaborated by Arora, Garg, & Gupta (2020).

Among various CNN architectures, ResNet50, a Residual Network with 50 layers, stands out with its 23 million trainable parameters, facilitating complex image classification tasks by capturing and reutilizing critical features. Moving beyond theoretical aspects, this article delves into the practical application of constructing a convolutional neural network for mask detection. Using a dataset from Kaggle, the article discusses methodologies for preprocessing and transforming images into vector data, leveraging ResNet50's capabilities for predicting mask-wearing status (Tian & Chen, 2019). Further, it explores the use of the confusion matrix as an essential tool for evaluating classification model performance, examining its components, calculations, and relevance in assessing precision, recall, accuracy, and other vital metrics.

2.1 Neural Networks

Artificial Neural Networks (ANNs) consist of numerous neurons that collectively enhance processing performance (O'Shea & Nash, 2015). Convolutional Neural Networks (CNNs) represent an advanced form of ANNs, featuring neurons that self-optimize through learning. These neurons in CNNs conduct vectorial mathematical operations regulated by non-linear functions, forming the computational foundation for numerous ANNs (FAIZAH et. al, 2021; SURESH, PALANGAPPA, & BHUVAN, 2021).

CNNs are specifically engineered for optimal image data processing. In this architecture, neurons are organized in three dimensions – height, width, and depth, where depth corresponds to the class variety in the dataset. In a departure from standard ANNs, neurons within a layer connect only to a specific region of the preceding layer. CNNs comprise convoluted layers, clustering layers, and fully connected layers (O'shea & Nash, 2015).

ResNet, an acronym for Residual Network, is a specific type of CNN introduced by He et al. (2016). ResNet50, a variant of this network, encompasses 48 convolution layers, a Max Pool, and an Average Pool layer, totaling more than 23 million trainable parameters, making it suitable for detailed classifications (He et al., 2016). ResNet50's ability to capture and repurpose key image features for smaller datasets significantly reduces data training time. In this model, images are resized to 224x224 pixels with three color channels, facilitating iterative use in various stages of the ResNet50 model (Islam, Tasnin & Baek et al., 2020).

2.2 Prediction model

The dataset employed in this study is sourced from Kaggle, a prominent platform hosting a myriad of datasets. The specific dataset utilized comprises images of individuals with and without face masks, totaling 1651 images divided into two distinct classes. Class 0, representing individuals wearing masks, includes 826 images, while Class 1, depicting individuals without masks, comprises 825 images. Following the dataset upload to Google Drive, the images were partitioned into training and testing sets. The training set included 658 images for Class 0 and 657 for Class 1. The testing set comprised 168 images from each class.

The computational approach involved the deployment of a Convolutional Neural Network (CNN), a critical component in the realm of computer vision. CNNs are renowned for their efficiency in pattern recognition, attributed to their reduced computational demand and proficiency in extracting data from images. Convolution, a linear operation in this context, combines two functions to produce a third entity, commonly referred to as a feature map. In image processing, this is analogous to the application of a filter/kernel transforming the input image through matrix multiplication, as discussed in Dekhtiar et al. (2018) and Huang et al. (2016).

The initial phase of model construction entailed importing essential libraries, including Keras and its derivatives. Subsequent steps involved incorporating the images and employing TensorFlow for resizing, modifying, and converting these images into vector data. The ImageDataGenerator library, instrumental in augmenting the number of images within the CNN, was utilized in the model development process. Post image generation, the trained model was loaded, and the classification of images was executed using ResNet50, a sophisticated CNN model. The primary function of ResNet50 in this context is to process input data (images) and make predictions regarding the presence or absence of face masks.

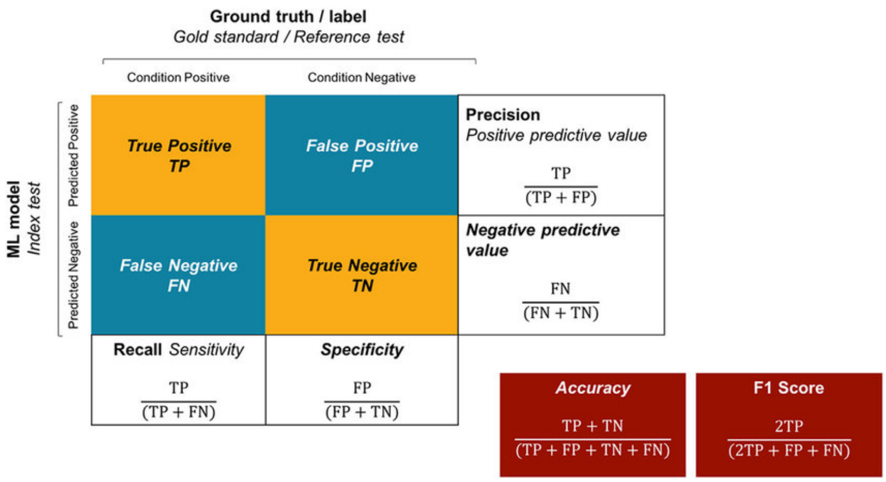
2.3 Confusion matrix

As defined by Sokolova & Lapalme (2009), a confusion matrix, or a contingency table, is an evaluative tool employed in supervised machine learning to assess the performance of classification models. Particularly pertinent in binary classification scenarios, it delineates the accuracy of predictions against actual labels, facilitating the analysis and refinement of the model's performance.

The confusion matrix, depicted in Figure 1, is structured as a rectangular table organizing a model's predictions relative to actual data labels. It comprises four quadrants representing various prediction outcomes. True Positives (TP) are instances where the model accurately identifies the positive class. True Negatives (TN) reflect correct predictions of the negative class. False Positives (FP) occur when the model inaccurately predicts the positive class, and False Negatives (FN) arise when it incorrectly predicts the negative class. This matrix serves as a crucial instrument in understanding and enhancing the precision, recall, accuracy, and other vital metrics of the model (Faes et al, 2009).

Figure 1

Overview of the confusion matrix.



Source: Faes et al (2020).

The confusion matrix serves as a comprehensive tool for evaluating the performance of a classification model, providing insights into specific error types made by the model. It facilitates the calculation of various key metrics such as precision, sensitivity (also known as recall), and specificity, among others. Figure 1 in the study highlights the terminological differences between machine learning (presented in bold) and classic statistical approaches (in italic type), including areas of overlap (denoted in bold and italic).

The application of the confusion matrix extends to assessing the quality of digital image classification. It achieves this by correlating information from reference data with the classified outcomes (positive or negative). This evaluation can also be conducted by juxtaposing training samples with the classified data. This latter approach is particularly prevalent in scientific research, especially in scenarios where reference data are absent, thus precluding the first method. The confusion matrix offers a comprehensive and effective means of presenting classification results, either in tabular or graphical formats.

The efficacy of this tool is particularly evident in binary classification tasks, where it aids in the calculation of accuracy, recall, precision, and the F1-score.

Accuracy is recognized as one of the most fundamental and crucial metrics, quantifying the percentage of correct predictions made by the model. Recall, or sensitivity, assesses the model's proficiency in correctly identifying positive results. Precision measures the proportion of true positives relative to the total number of positive predictions made by the model. The F1-score, or F1 measure, represents a harmonic mean of precision and recall, offering a balanced metric that considers both false positives and false negatives (Kohavi & Provost, 1998).

3. METHODOLOGICAL PROCEDURES

According to Bradski & Kaehler (2008), computer vision transforms data captured by cameras into decisions or representations for specified purposes. The input data may include contextual information such as the camera's location or its distance within a scene. This process enables decisions like identifying specific objects, including the presence of individuals with or without face masks in a scene.

The primary development environment utilized for this research was Google Collaboratory (Google Colab), a free Jupyter notebook environment running in the cloud. Google Colab facilitates code writing, execution, analysis sharing, and access to robust computing resources, such as the Keras and TensorFlow libraries, without any setup requirement.

The study employed neural networks, specifically ResNet50, to evaluate the results using a confusion matrix as the primary metric. The OpenCV library, initially developed by Intel in 2000, was used for digital image manipulation, thereby democratizing computer vision access across various domains. Additionally, Keras, an interface for streamlined deep learning model implementation, and TensorFlow, a symbolic mathematics library widely utilized in machine learning applications, were integral to the algorithm development.

The Python script developed for this study constructs and evaluates a convolutional neural network (CNN) model using TensorFlow and Keras, targeting image processing tasks. The script begins with importing necessary libraries, including TensorFlow, Keras, NumPy, pandas, seaborn, and matplotlib. ImageDataGenerator was utilized for image pre-processing and data augmentation. The model is built as a sequential arrangement of layers, such as Conv2D, MaxPooling2D, Flatten, Dense, Dropout, and BatchNormalization. Evaluation metrics like confusion matrix and classification report are employed, along with functions for saving the trained model.

The training and evaluation of the model utilized a total of 1315 images, divided into 979 for training and 336 for testing. The script defines image generators for both sets using ImageDataGenerator, applying data augmentation techniques for training to enhance data variety and mitigate overfitting. For testing, a simpler generator is used, both incorporating specific pre-processing functions for ResNet50.

ResNet50 forms the base model, loaded with pre-trained weights from ImageNet but excluding the top classification layer. New dense layers are added for the specific task of classifying mask-wearing status, culminating in a softmax activation layer for binary classification. The training process involves freezing the base ResNet50 model up to layer 175, training only the newly added layers. The model is compiled with categorical_crossentropy as the loss function and accuracy as the metric, undergoing training for a designated number of epochs and steps per epoch.

The script also provides a quantitative and visual analysis of the model's performance over time, utilizing training and testing metrics stored in the history object. Metrics such as mean and standard deviation of accuracy for training and testing data are calculated, and loss and accuracy metrics are graphically represented using matplotlib. This analysis helps identify learning patterns and potential issues like overfitting or underfitting.

The evaluation of the machine learning model focuses on accuracy and the confusion matrix for testing data. Accuracy is assessed using the accuracy_score function from sklearn.metrics, and the confusion_matrix function generates confusion matrices for the test data. The seaborn library is employed to create heatmaps, offering a visual representation of the actual versus predicted classes. The classification_report provides detailed class-by-class analysis, including precision, recall, and F1 score metrics.

The mentioned libraries and software packages are accessible at their respective online resources: TensorFlow (https://www.tensorflow.org), Keras (https://keras.io), NumPy (https://numpy.org), Scikit-learn (https://scikit-learn.org), Pandas (https://pandas.pydata.org), Seaborn (https://seaborn.pydata.org), and Matplotlib (https://matplotlib.org). These platforms offer extensive documentation, tutorials, and user guides for these tools.

4. RESULTS AND DISCUSSION

In this study, deep learning and computer vision techniques were synergistically employed to develop a system for monitoring compliance with mask-wearing guidelines in public spaces. This system automates the process of determining whether individuals are wearing masks, addressing a crucial public health measure during the COVID-19 pandemic.

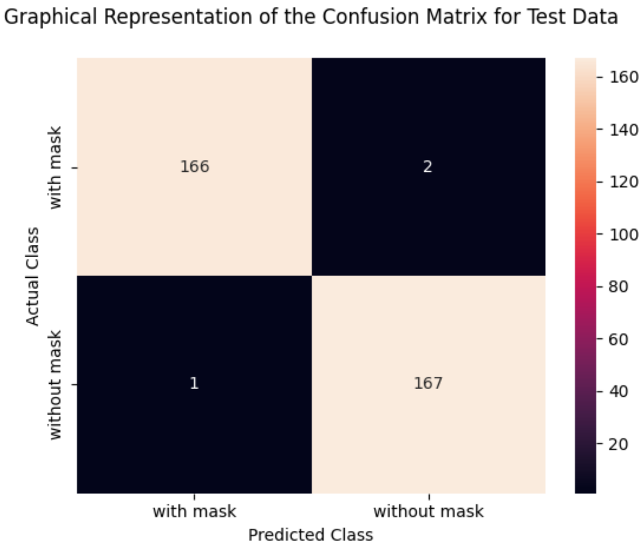
The study involved the collection of images depicting individuals both with and without masks under varying lighting conditions. These images served as the dataset for training and testing a convolutional neural network (CNN), which was the primary tool for image classification. The CNN models, grounded in artificial intelligence (AI), were meticulously designed and programmed to transform these images into statistical data. The models underwent rigorous training to accurately classify images into two categories: individuals wearing masks (class 0) and those without masks (class 1).

Binary classification studies such as this often utilize a confusion matrix to tabulate the accuracy of the model's predictions relative to actual data. In total, 1,315 images were utilized for training the model. The effectiveness of the model was evaluated using a confusion matrix for the testing procedure, as depicted in Figure 2. This matrix clearly delineates the model’s accuracy, with successful predictions represented on the main diagonal and errors indicated on the secondary diagonal.

For the validation of the trained model, a set of 336 images was used. The results, presented in Figure 2, demonstrate that the model correctly identified 166 images as true positives (individuals wearing masks) and 167 images as true negatives (individuals not wearing masks). The secondary diagonal, representing prediction errors, showed 2 false negatives (masked individuals incorrectly identified as not wearing masks) and 1 false positive (an unmasked individual incorrectly identified as wearing a mask).

This study's findings highlight the potential of AI-driven CNN models in automating and enhancing public health monitoring systems, particularly in the context of the ongoing global health crisis.

Figure 2
Test confusion matrix.



In the context of this study, the term 'accuracy' is defined as the degree of exactness of a measurement, where an accurate measurement closely approximates the true value. For the specific application of this study, accuracy is considered in the realm of a classifier designed to differentiate between images of individuals wearing masks (class 0) and those without masks (class 1). The classifier's database, comprising 1,651 images, served as the basis for both training and testing phases.

Precision, in this analytical framework, is conceptualized as the degree of consistency and repeatability in measurements. A more precise measurement displays minimal variation between individual measured values. The metric of recall, or sensitivity, is employed to quantify the proportion of accurately identified positive predictions. Recall becomes particularly significant when the objective is to minimize false negatives, as a higher recall rate indicates a reduced incidence of false negative results. The F1-score, as defined by Kohavi & Provost (1998), represents the harmonic mean of precision and recall. This metric assumes a value akin to the arithmetic mean when precision and recall are closely aligned.

Regarding the test data, the accuracy metric, which reflects the degree of correspondence with the true value, averaged approximately 99% across a sample of 336 images. Table 1 delineates the specific values for precision, recall, and F1-score, providing a detailed breakdown of these metrics for the test images. This breakdown includes data for both class 0 (individuals with masks) and class 1 (individuals without masks).

Table 1
Table of test metrics.

	PRECISION	RECALL	F1-SCORE	IMAGES
Class 0 (with mask)	0.99	0.99	0.99	168
Class 1 (without mask)	0.99	0.99	0.99	168

In Figure 3a, the performance of the algorithm is delineated in terms of accuracy metrics over time. This figure elucidates the temporal evolution of the algorithm's accuracy throughout its training and testing phases.

Notably, the training phase is represented by a blue curve, characterized by minimal fluctuations and an overall trend towards stabilization. In contrast, the testing phase, depicted by an orange curve, demonstrates slight variations but exhibits a rapid convergence to an accuracy exceeding 95% post the 10th epoch.

Conversely, Figure 3b presents the error metrics as a function of seasonal variations. Within this figure, the training phase is illustrated by a blue curve, indicating a downward trajectory approaching zero. The testing phase, however, represented by an orange curve, displays a series of fluctuating peaks occurring at distinct intervals. This phase initiates with minimal values, spanning epochs 0 to 60, and culminates in a maximal error rate at the 75th epoch.

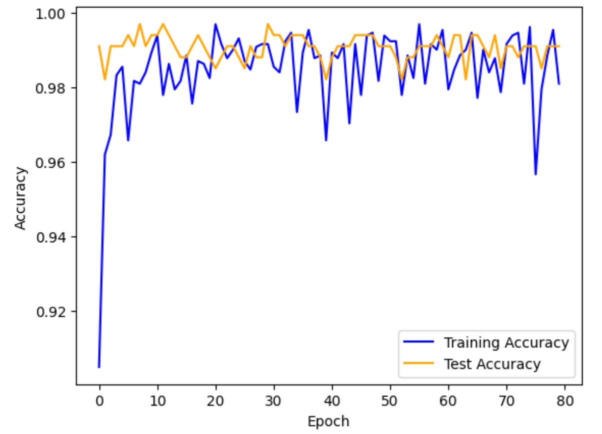


Figure 3a: Graph depicting variation in accuracy metrics across different seasons.

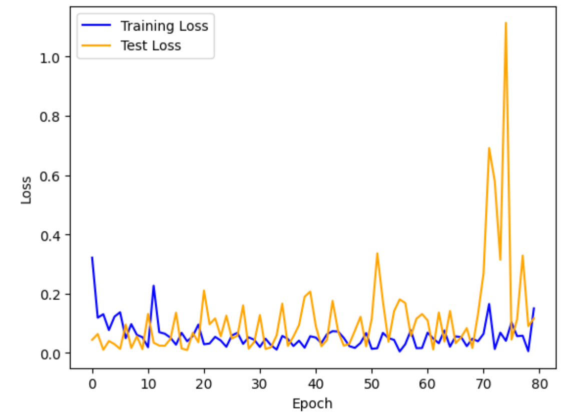


Figure 3b: Graphical representation of loss (error) variation as a function of seasonal factors.

This study demonstrates the capability of computer vision, coupled with a classification algorithm, to predict the presence or absence of face masks in input images. This application of computer vision technology serves as a monitoring tool for adherence to COVID-19 preventative guidelines, illustrating its potential in mitigating the health crisis.

Computer vision, a domain within computer science and artificial intelligence, is dedicated to analyzing, interpreting, and extracting meaningful information from images for decision-making or generating valuable data for various applications. In the context of the COVID-19 pandemic, computer vision emerges as a pivotal technology for monitoring compliance with health guidelines, thereby contributing to public health efforts during these challenging times. The adopted model holds significant potential for academic and professional settings, offering a creative and efficient technological solution.

Comparative analysis with related studies, such as 'Face Mask Detection using Optimistic Convolutional Neural Network' (SURESH, PALANGAPPA & BHUVAN, 2021) and 'Implementation of the Convolutional Neural Network Method to Detect the Use of Masks,' (FAIZAH, 2021) reveals varying yet effective applications of Convolutional Neural Networks (CNNs) and OpenCV in the realm of facial mask detection. The first study integrates MobileNet and OpenCV with advanced deep learning algorithms and geometric techniques, achieving an impressive training accuracy of 99.9% and 100% in testing.

This outcome underscores the model's robustness across diverse datasets. The second study employs OpenCV's Haar Cascade Classifier in conjunction with a CNN model, yielding training and testing accuracies of 98.86% and 96.19%, respectively. The current study utilizes the ResNet50 neural network architecture along with OpenCV for digital image processing, attaining an average test precision of approximately 99%, with comparably high precision, recall, and F1-score metrics.

The findings of this study, showcasing a 99% accuracy rate in image prediction tasks using deep learning for face mask detection, provide an important benchmark in the realm of computer vision, particularly in the context of the COVID-19 pandemic. To contextualize these results, a comparison with similar studies is essential.

Militante & Dionisio (2020) conducted a study using deep learning for facial recognition and mask detection, with a dataset of 25,000 images at a 224x224 pixel resolution. Their model achieved an accuracy rate of 96%. While this is slightly lower than the 99% accuracy reported in our study, it still represents a high level of precision in similar tasks, indicating the effectiveness of deep learning techniques in this field.

Sethi, Kathuria & Kaushik (2021) explored the integration of various models, including ResNet50, with their proposed technique, achieving an impressive 98.2% accuracy. This result is very close to the 99% accuracy of our study, further validating the potential of deep learning algorithms in efficiently recognizing face masks.

In addition to the studies, the researchers conducted by Kodali & Dhanekula (2021) and Suryadevara (2020, 2021) also demonstrates a notable degree of precision in the field of computer vision using deep learning techniques. These studies successfully achieved an accuracy rate exceeding 95% in their image prediction tasks. This high level of precision further corroborates the efficacy of deep learning algorithms in accurately performing complex tasks such as image recognition and analysis, particularly in applications requiring detailed and precise outcomes. Such consistent high performance across diverse studies highlights the robustness and reliability of these techniques in advancing computational image processing and recognition.

These comparisons suggest that high accuracy rates are common in studies leveraging deep learning for face mask detection. The slight variations in accuracy can be attributed to differences in datasets, algorithmic implementations, and model architectures. The consistent achievement of high accuracy across different studies underscores the reliability and effectiveness of deep learning techniques in addressing public health challenges, such as the detection of face mask usage during the COVID-19 pandemic.

These multifaceted methodologies underscore the versatility and effectiveness of CNNs and related techniques in the realm of public health, especially in the context of the COVID-19 crisis response. Despite their individual distinctiveness, these varied approaches cohesively illuminate the pivotal role that computer vision and deep learning play in tackling current health challenges. This convergence of methods not only demonstrates the applicability of these technologies in diverse scenarios but also reinforces their significance as integral tools in the evolving landscape of global health crises management.

5. CONCLUSION

This study demonstrates that deploying deep learning algorithms and computer vision techniques for monitoring face mask usage in public places offers a promising strategy for aiding in the prevention and control of COVID-19. Additionally, the model developed herein holds potential for adaptation to detect other essential protective accessories, providing substantial benefits in academic and professional settings. The application of artificial intelligence for mask detection and monitoring emerges as a significant tool in public health efforts to curb the spread of COVID-19 and other infectious diseases.

However, it is crucial to acknowledge certain limitations of this study. The dataset employed to train the model was relatively limited in size, possibly constraining its capacity to fully represent the diversity of real-world mask-wearing scenarios.

Furthermore, the model's focus was solely on the binary detection of mask presence or absence, without considering factors such as mask quality or correct fit.

As the current model represents a theoretical framework, its practical implementation and validation in real-world settings remain necessary. Despite these limitations, the approach outlined in this study is a noteworthy advancement in the use of computer vision and deep learning for public mask usage monitoring. Considering the results and limitations identified, the following suggestions are proposed for future research:

a) Enlargement and Diversification of the Dataset: To enhance the model's generalizability across varied scenarios, a larger and more diverse dataset encompassing various mask types, lighting conditions, and camera angles is recommended.

b) Development of Models for Quality and Fit Assessment: Future models should not only detect masks but also evaluate their quality and whether they are worn correctly to ensure effective protection.

c) Utilization of Transfer Learning Techniques: Leveraging pre-trained models on related tasks, such as face detection or object recognition, and adapting them for mask detection could prove more efficient than training models from scratch.

d) Extension to Other Protective Accessories: Investigating the detection of additional protective gear, like gloves or face shields, is a worthwhile avenue, given their importance in infectious disease prevention.

e) Real-world Validation Studies: Conducting validation studies in environments where mask detection is critical, such as airports, hospitals, or workplaces, would provide valuable insights into the model's efficacy and accuracy in practical applications.

These recommendations aim to guide future research in enhancing computer vision algorithms for mask detection and monitoring, thereby contributing to public health safety measures.

REFERENCES

Apostolopoulos, Ioannis D. & Mpesiana, Tzani A. (2020). Covid-19: automatic detection from x-ray images utilizing transfer learning with convolutional neural networks. **Physical and Engineering Sciences in Medicine**, v. 43, p. 635-640. https://doi.org/10.1007/s13246-020-00865-4

Arora, Divya, Garg, Mehak & Gupta, Megha. (2020). Diving deep in deep convolutional neural network. In: **2nd International Conference on Advances in Computing, Communication Control and Networking** (ICACCCN). IEEE, p. 749-751. https://doi.org/10.1109/ICACCCN51052.2020.9362907.

Bradski, Gary & Kaehler, Adrian. (2008). Learning OpenCV: **Computer vision with the OpenCV library**. O'Reilly Media, Inc. ISBN: 978-0-596-51613-0

Dekhtiar, Jonathan et al. (2018). Deep learning for big data applications in CAD and PLM–Research review, opportunities, and case study. **Computers in Industry**, v. 100, p. 227-243. https://doi.org/10.1016/j.compind.2018.04.005

Faes, Livia et al. (2020). A clinician's guide to artificial intelligence: how to critically appraise machine learning studies. **Translational Vision Science & Tecnology**, v. 9, n. 2, p. 7-7. https://doi.org/10.1167/tvst.9.2.7

Faizah, Arbiati et al. (2021). Implementation of the convolutional neural network method to detect the use of masks. **International Journal of Informatics and Information Systems**, v. 4, n. 1, p. 30-37. https://doi.org/10.47738/ijiis.v4i1.75

He, Kaiming et al. (2016). Deep residual learning for image recognition. In: **Proceedings of the IEEE conference on computer vision and pattern recognition**. p. 770-778. http://openaccess.thecvf.com/content_cvpr_2016/html/He_Deep_Residual_Learning_CVPR_2016_paper.html

Huang, G. et al. (2016) Densely connected convolutional networks. **Arxiv website**. arxiv.org/abs/1608.06993. Published August, v. 24. http://openaccess.thecvf.com/content_cvpr_2017/html/Huang_Densely_Connected_Convolutional_CVPR_2017_paper.html

Islam, Md Mahbubul, Tasnim, Nusrat, & Baek, Joong-Hwan. (2020). Human gender classification using transfer learning via Pareto frontier CNN networks. **Inventions**, v. 5, n. 2, p. 16. https://doi.org/10.3390/inventions5020016

Kodali, Ravi Kishore, & Dhanekula, Rekha. (2021). Face mask detection using deep learning. In: **International Conference on Computer Communication and Informatics** (ICCCI). IEEE, p. 1-5. https://doi.org/10.1109/ICCCI50826.2021.9402670

Kohavi, Ron & Provost, Foster. (1998). Glossary of Terms: special Issue on Applications of Machine Learning and the Knowledge Discovery Process. **Journal of Machine Learning**, 30 (2–3) (1998), pp. 271-274. http://ai.stanford.edu/~ronnyk/glossary.html

Militante, Sammy V. & Dionisio, Nanette V. (2020). Real-time facemask recognition with alarm system using deep learning. In: **11th IEEE Control and System Graduate Research Colloquium** (ICSGRC). IEEE, p. 106-110. https://doi.org/10.1109/ICSGRC49013.2020.9232610

O'shea, Keiron & Nash, Ryan. (2015). An introduction to convolutional neural networks. **arXiv preprint arXiv:1511.08458**. https://doi.org/10.48550/arXiv.1511.08458

Sethi, Shilpa; Kathuria, Mamta & Kaushik, Trilok. (2021). Face mask detection using deep learning: An approach to reduce risk of Coronavirus spread. **Journal of biomedical informatics**, v. 120, p. 103848. https://doi.org/10.1016/j.jbi.2021.103848

Sokolova, Marina & Lapalme, Guy. (2009). A systematic analysis of performance measures for classification tasks. **Information Processing & Management**, v. 45, n. 4, p. 427-437. https://doi.org/10.1016/j.ipm.2009.03.002

Suresh, K., Palangappa, M. B. & Bhuvan, S. (2021). Face mask detection by using optimistic convolutional neural network. In: **6th International Conference on Inventive Computation Technologies (ICICT)**. IEEE. p. 1084-1089. https://doi.org/10.1109/ICICT50816.2021.9358653

Suryadevara, Chaitanya Krishna. (2020). Real-time face mask detection with computer vision and deep learning. **International Journal of Innovations in Engineering Research and Technology**, v. 7, n. 12. https://ssrn.com/abstract=4591986

Suryadevara, Chaitanya Krishna. (2021). Enhancing safety: face mask detection using computer vision and deep learning. **International Journal of Innovations in Engineering Research and Technology**, v. 8, n. 8. https://ssrn.com/abstract=4592001

Tian, Xiao & Chen, Chao. (2019). Modulation pattern recognition based on Resnet50 neural network. In: **2019 IEEE 2nd International Conference on Information Communication and Signal Processing** (ICICSP). IEEE. p. 34-38. https://doi.org/10.1109/ICICSP48821.2019.8958555

Yadav, Shashi et al. (2020). Deep learning based safe social distancing and face mask detection in public areas for covid-19 safety guidelines adherence. **International Journal for Research in Applied Science and Engineering Technology**, v. 8, n. 7, p. 1368-1375. https://doi.org/10.22214/ijraset.2020.30560