IAES International Journal of Artificial Intelligence (IJ-AI)
Vol. 13, No. 3, September 2024, pp. 3703~3714
ISSN: 2252-8938, DOI: 10.11591/ijai.v13.i3.pp3703-3714

Artificial intelligence ethics: ethical consideration and regulations from theory to practice

Shurooq Mnawer Ibrahim[1], Mohammad Aref Alshraideh[2,3], Martin Leiner[4], Iyad Muhsen Aldajani[4], Bettaz Ouarda[3]

[1]Department of Computer Science, The University of Jordan, Amman, Jordan
[2]Department of Artificial Intelligence, The University of Jordan, Amman, Jordan
[3]Information Technology College, Lusail University, Lusail, Qatar
[4]Jena Center for Reconciliation Studies, Friedrich Schiller University, Jena, Germany

Article Info

Article history:

Received Nov 3, 2023
Revised Feb 14, 2024
Accepted Mar 6, 2024

Keywords:

Algorithmic bias
Artificial intelligence governance
Bias mitigation techniques
Ethical frameworks
Responsible artificial intelligence

ABSTRACT

The advancement of artificial intelligence (AI) has led to its widespread use in sectors such as finance, healthcare, military, and employment in developed countries. However, this reliance has raised concerns about AI governance, particularly regarding algorithmic biases based on skin color, gender, race, and age. Consequently, many countries have introduced regulations and ethical frameworks to address these issues. The Ministry of Digital Economy and Entrepreneurship in Jordan has included AI in its 2022 plan, signaling significant progress. The integration of AI in education programs underscores this commitment. However, addressing AI's potential negative impacts is essential. We propose ethical considerations and regulations for AI to complement Jordan's initiatives. Our research aims to promote responsible AI usage by developing ethical guidelines in Jordan. It presents techniques to identify and mitigate biases related to skin color, gender, and age in AI outputs and datasets. The research includes extensive testing on datasets, analyzing approximately 100 images, and revealing notable error rates, including a 16% error rate in detecting skin color, a 4% error rate in seeing white faces, and a 6% error rate in identifying females over men. Therefore, ethical considerations and regulations for AI applications in Jordan must be implemented.

Corresponding Author:

Mohammad Alshraideh
Department of Artificial Intelligence, The University of Jordan
Amman, Jordan
Email: mshridah@ju.edu.jo

1. INTRODUCTION

Human morality has developed over thousands of years through various experiences, including conflicts, alliances, marriages, and contributions from poets, writers, philosophers, and leaders. These experiences have led to rules and foundations that guide human behavior with respect and fairness. On the other hand, artificial intelligence (AI) involves creating entities like machines, robots, or applications to control and harness their abilities to enhance human lives. AI is used in various domains, from human-computer interfaces to cyber-physical systems such as the Internet of Things and robotics.

During a UNESCO conference held in Paris in 2021 and documented in 2022 [1], AI was defined as a system capable of processing data and information in a way that resembles human intelligence. These systems use algorithms and models to perform cognitive tasks, learn from data, make decisions and predictions, and plan actions. They include learning, reasoning, perception, and control.

Despite its progress and benefits, AI has brought about significant impacts and challenges, as outlined by the European Parliamentary Research Service (ERPS) [2]. These impacts and challenges can be grouped into several categories:

- Social influences: AI affects the labor market, the economy, and demographic groups, raising concerns about the concentration of power and authority in certain areas. It also raises concerns about privacy, human rights, dignity, and bias.
- Psychological influences: AI's interactions with humans, especially robots, can psychologically impact people's well-being.
- Financial influences: AI can impact the economy and financial systems, necessitating measures to prevent manipulation and collusion, and enhance accountability.
- Legal influences: AI systems can disrupt the global legal system by increasing the risks of criminal activities, including trafficking in human organs, drugs, and prostitution. This necessitates regulatory controls.
- Environmental influences: AI can offer potential benefits in addressing environmental challenges like waste management and energy efficiency.

AI encompasses robots that interact with real-world environments, including city streets for autonomous vehicles, human settings for social robots, and various other contexts. These interactions have necessitated the development of AI ethics, which emphasizes fairness and transparency. Machine learning is integral to AI, allowing robots to learn and adapt to their environments through supervised and unsupervised learning techniques.

According to the transportation research board, AI has positive and negative impacts across fields such as medicine, engineering, law, economics, and security [3]. The EPRS [2] defines ethics as the principles governing conduct and behavior, emphasizing respect for others. Various philosophers, including Kant, have offered different definitions and principles for ethics, generally agreeing that it is a system of moral rules guiding individual and group behavior. Kidder [3] described ethics as studying the ideal human personality or ethical duty.f

AI ethics, as outlined by EPRS [2], refers to a system of ethical principles and strategies for the responsible use of AI. Companies have already started creating rules for AI, acknowledging its significant role in our daily lives. AI ethics focuses on the actions of manufacturers, operators, and developers to mitigate the potential harm that AI can inflict on society due to improper application, unethical design, or misuse.

The scope of AI ethics encompasses contemporary concerns such as bias in AI systems, data privacy, and the potential short- and long-term consequences, including the impact of AI and robotics on employment and the accessibility of AI systems with capabilities equal to or surpassing human abilities. AI significantly influences various sectors, including healthcare, finance, transportation, and entertainment. Therefore, establishing a set of ethical principles and guidelines is crucial to ensure that the development and use of AI are responsible, fair, and safe. Given AI's profound effects on careers, health, and economies, including the potential for job displacement, ethical considerations are imperative, as highlighted by research conducted at the University of Jordan.

Establishing ethical considerations and legal regulations for using AI in technological applications and scientific research is crucial. This is particularly important in Jordan, where AI is integrated into legal and educational systems. Recommendations must be created for its use to ensure Ethical AI, addressing challenges and future work in conducting ethical and legal considerations on integrated internet research methodologies with AI applications. To achieve this, AI algorithms were utilized on various datasets related to skin color, age, and gender ratios to provide recommendations for developing AI applications in research. The article is structured as follows: section 2 shows the evolution of the literature review and background, section 3 introduces the methodology, section 4 presents the discussion and results, section 5 provides recommendations and discusses the challenges, and finally, section 6 concludes the article findings and related future work.

2. LITERATURE REVIEW AND BACKGROUND

AI Ethics in Jordan: while the passage provides an overview of Jordan's efforts to promote AI, it is essential to emphasize the significance of ethical considerations in Jordan's AI journey. Ethical considerations become paramount as AI technologies advance and integrate into various aspects of society. Jordan's focus on AI education and initiatives like the "AI citizenship in Jordan" platform demonstrates a commitment to harnessing AI's potential for the benefit of its citizens [4].

AI and COVID-19: according to Shaikh *et al.* [5], the COVID-19 pandemic accelerated the adoption of AI in various countries, including Germany. The passage highlights the role of AI-powered apps like the Corona-Warn app in tracking and managing the pandemic. However, it also underscores the importance of

ethical guidelines and transparency in AI applications during a public health crisis. Privacy concerns and algorithm bias need to be addressed to maintain public trust.

Algorithm bias: according to Floridi and Cowls [6], algorithm bias is a pervasive issue in AI development. The passage discusses how bias can emerge from inaccurate or incomplete data, leading to discriminatory outcomes. It is crucial to delve deeper into this topic, as algorithmic bias can have far-reaching consequences, affecting criminal justice, hiring practices, and financial lending. Addressing bias in AI algorithms requires rigorous data preprocessing and continuous monitoring to ensure fairness and transparency.

Global collaboration: according to Hooker and Kim [7], the passage highlights the need for international collaboration on AI ethics and regulations. The European Union efforts to develop ethical guidelines and the involvement of organizations like UNESCO indicate a growing awareness of the need for a global framework. As AI technologies transcend borders, international cooperation becomes essential to harmonize standards and ensure responsible AI development worldwide.

Education and awareness: according to Alshamaila *et al*. [8], the passage underscores the importance of AI education and awareness initiatives. Jordan's introduction of AI programs in universities and the focus on ethical AI education in EU guidelines demonstrate a recognition of the need to prepare future generations for an AI-driven world. Building a workforce that understands the moral implications of AI is crucial for responsible development and usage.

In the realm of AI ethics, various studies and organizations have laid the groundwork for ethical AI development and deployment. Here is a summary of these critical studies and initiatives. Several researchers [3], [9], [10] developed a framework to analyze and compare different sets of AI ethical principles proposed by various entities. They identified five core principles: benevolence, non-abuse, independence, justice, and interpretability. These principles are a foundation for understanding how AI should function and guide its development and use. Their analysis found that these principles align closely with those proposed by IEEE and Montreal, emphasizing the importance of creating AI that benefits humanity while promoting well-being.

According to Kovač *et al.* [11], the IBM Institute proposed a comprehensive project to establish ethics in AI and raise awareness about the need for ethical considerations in AI development. They highlighted five ethical areas: accountability, value alignment, explainability, fairness, and user data rights. IBM stressed the role of designers and developers in minimizing biases and safeguarding human rights in the context of AI technologies.

Levi and Hassncer [12] develops a deep learning model to predict student performance, tackling class imbalance. It uses a dataset from the University of Jordan, exploring various features. Results indicate strong performance, offering insights for education management. Future research can build upon these methods for further advancements. UNESCO [13] explored instances of bias in machine learning algorithms, particularly in areas like employment, housing, and insurance. The author emphasized the challenges posed by implicit biases in AI and suggested measures to detect and mitigate these biases. Stricter public policies and increased efforts to address algorithmic bias were recommended. Roberts *et al.* [14] proposed an approach to AI ethics that incorporates both human intuition and formal logic. They introduced moral principles based on quantitative reasoning, emphasizing the importance of transparency in AI systems. The concept of AI ethics is seen to create transparent AI systems that provide explanations for their actions.

According to Southgate *et al.* [15], Oxford University founded the Institute for Ethics in AI, recognizing the urgent need to address the ethical challenges posed by AI technologies. The institute aims to embed ethics in AI education and research, making AI ethics a global field. They prioritize interdisciplinary collaboration and have initiated seminars and discussions to engage various stakeholders in AI ethics. Southgate [16] advocated for a theoretical approach to address AI challenges, especially when dealing with complex definitions like privacy and fairness. They proposed setting ethical goals early in AI development and introduced scientific methods to achieve justice and fairness in AI designs. In [8], [17], a framework for introducing ethical AI into educational systems, including higfher education. Their framework combines technology and human rights, systematically identifying and responding to ethical challenges in AI education. It emphasizes the importance of aligning AI education with fundamental human rights principles. Southgate [16] emphasize the crucial principle of data privacy in AI and machine learning. They provide several compelling examples to underscore the significance of safeguarding data privacy.

- GIC insurance commission example**:** In the 1990s [2], the GIC insurance commission in Massachusetts shared hospital visit data of government employees with academic researchers, with personally identifiable information removed. However, Latanya Sweeney, a doctoral student at MIT, demonstrated that it was possible to re-identify individuals from seemingly anonymized data. She used voter rolls to identify the governor's medical records, highlighting the potential re-identification risks even with limited information. Sweeney's research revealed that 87 percent of the US population could be identified using just three data points.
- Netflix prize: Netflix hosted the Netflix Prize Competition in 2006 to develop a movie recommendation

algorithm. They released a large movie ratings dataset to facilitate this while removing user IDs. However, Arvind Narayanan and his advisor showed that an attacker could re-identify individuals from the supposedly anonymized records even with limited background information. This raised concerns about the privacy of Netflix users.

Southgate [16] emphasize that research studies using data can potentially inadvertently harm the individuals whose data is used when published. For example, a survey conducted in the UK in 1951 linking smoking to lung cancer resulted in potential consequences for one of the doctors whose data was used. This example highlights the challenges in maintaining data privacy in AI and machine learning, underscoring the need for stringent privacy protection measures. The rapid development of AI technology brings significant benefits but raises concerns about potential biases, ethical considerations, and their impact on various aspects of society. International efforts have been made to address these concerns. Jobin *et al*. [9] identified six widely accepted ethical principles in AI: non-offence, transparency, fairness, justice, privacy, and responsibility. The United Nations has recognized over 160 international, national, and organizational groups focusing on AI ethics and governance, emphasizing the need for a unified platform to coordinate these efforts.

Stanford's Institute for Human-Centered Artificial Intelligence (HAI) [18] discuss the increasing ethical challenges of AI technologies, such as privacy intrusion and discrimination. It explores efforts to address these issues, including developing AI principles, media coverage, research presentations, and university ethics courses. However, a lack of quantitative data makes measuring societal discussions' impact on technology development hard. Policymakers prioritize measurable data, so translating qualitative arguments into quantitative data is crucial for effective policymaking.

Esther and Tella [19] focuses on the ethical implications of AI in the United States, emphasizing the need for ethical guidelines throughout its development and use. It outlines eleven fundamental principles, including transparency, fairness, and privacy, to guide responsible AI implementation. The paper explores the impact of AI, particularly machine learning, and discusses approaches to implementing AI ethics to address growing concerns about associated risks.

Pekka *et al.* [20] explores the ethics of AI and the transparency challenge in implementing ethical AI practices. It highlights explainable artificial intelligence (XAI) as a solution and presents findings from a systematic mapping study (SMS). The study identifies research gaps and provides empirical insights, offering potential directions for further AI ethics.

Singh *et al.* [21] examines the integration of AI into Institutional Ethics Committee reviews of clinical research. It highlights AI's potential benefits, like increased productivity and standardized assessments, and challenges, like loss of human insight and ethical concerns. The study proposes a roadmap for future AI integration, emphasizing collaboration, ethical design, and stakeholder engagement. Overall, it underscores the need to carefully balance technological innovation and moral integrity in clinical research ethics.

Hallamaa and Kalliokoski [22] examines the need for sustainable design and development of AI and suggests that current AI ethics frameworks may lack methodological solidity. It proposes leveraging insights from bioethics to enhance the effectiveness of AI ethics. The article advocates incorporating tools from fields like systems theory, safety research, impact assessment, and theory of change to improve the quality of AI ethics and its influence on AI design and development practices.

While embracing AI's immense potential, we must acknowledge and responsibly address its inherent concerns, risks, and potential for misuse. This does not mean hindering innovation; it requires prioritizing ethical AI practices grounded in human rights, values, and principles. Doing so can foster responsible research and innovation and safeguard individual rights, freedoms, and dignity.

The global landscape of AI ethics and regulations is rapidly evolving, with various countries and regions taking strides to address the challenges and opportunities presented by this technology. However, the journey does not end here. Continued education, transparency, and international collaboration efforts are crucial to building a future where AI serves society while adhering to ethical principles and respecting human rights. Let us work together to shape a future where the benefits of AI are accessible to all without compromising our values and humanity.

3. PROPOSED METHODOLOGY

This section investigates bias in AI algorithms for skin tone, gender, and age. It details detection algorithms, exploring grayscale conversion and thresholding for skin color, and deep learning models with convolutional neural networks (CNN) for gender and age. Real-world case studies will later examine the societal impact of such biases.

3.1. Skin tone detection algorithm

The skin tone detection algorithm is designed to identify and quantify bias related to skin color in image-based AI algorithms. Here are the steps involved:

a) Convert the image to a grey level: the algorithm converts the input image to grayscale using the rgb2gray directive. This eliminates hue and saturation information while preserving luminance.
b) Detect and classify skin color: the algorithm analyzes the image's dimensions, precisely the width and height of the skin color region. It applies conditional rules based on Kovač *et al.* [11] to detect and classify skin color. These rules consider factors like red (R), green (G), and blue (B) channel values in the RGB color space, which vary based on lighting conditions. A nested for loop applies these rules to each pixel in the image. Pixels matching the skin color criteria are considered part of the skin region.
c) Skin color threshold: the algorithm employs two methods of binary quantization to determine whether the skin color is white or black. Grey thresholds are applied to the pixels within the skin color region and the pixels within the face area. The total threshold is calculated as the average of the skin color threshold and the face area threshold. The threshold value is used to distinguish between white and black skin. If the total threshold is more significant than the default threshold (based on Kovac *et al.*'s equations), the skin color is classified as white or black.
d) Percentage of detected faces: the algorithm calculates the percentage of faces detected in the images by dividing the number of detected faces by the total number of pictures and multiplying by 100.
e) Percentage of white and black faces: the algorithm calculates the Percentage of white and black faces detected among the total number of faces seen.

3.2. Analyzing gender and age detection in images

The following text describes an algorithm that detects gender and age in images and quantifies bias. The algorithm relies on deep learning models based on CNNs and is trained on a dataset provided by Levi and Hassncer [12]. The algorithm uses OpenCV, a versatile library for computer vision tasks, to implement gender and age detection. OpenCV's machine learning library (MLL), which includes statistical patterns and clustering algorithms to support gender and age prediction, is employed. The algorithm involves various preprocessing steps, thresholding, and leveraging pre-trained deep learning models to achieve its objectives. It utilizes a specific CNN architecture to identify faces, determine gender, and classify age into particular categories.

The CNN architecture used in this algorithm is relatively straightforward, consisting of the following layers, as described by Levi and Hassinger [12] in Figure 1:

a) Convolutional layer 1 (Conv1):
– This layer comprises 96 nodes.
– It uses a kernel size of 7.
b) Convolutional layer 2 (Conv2):
– The second layer consists of 256 nodes.
– It employs a kernel size of 5.
c) Convolutional layer 3 (Conv3):
– Layer 3 contains 384 nodes.
– It uses a kernel size of 3.

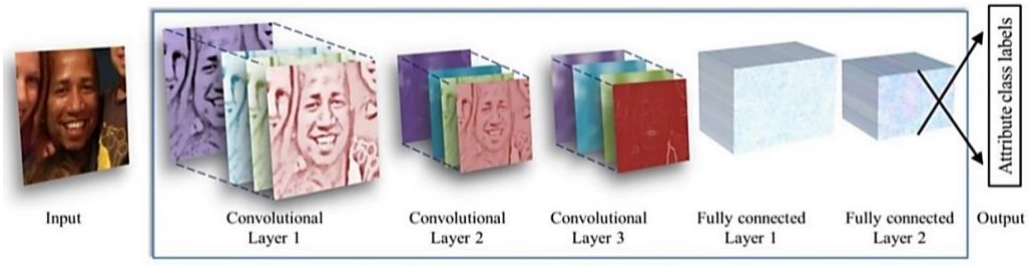

Figure 1. CNN architecture that we used [23]

The architecture includes two fully connected layers with 215 nodes each. In summary, this CNN architecture was designed to analyze facial features and extract data related to age and gender from input images. It consists of three convolutional layers and two fully connected layers. It ultimately leads to the final output layer for predicting gender and age.

Categories of gender and age: this section will explain how our algorithm detects gender and age. Our gender detection is a classification task with two categories: male and female. We formulate this as a gender prediction problem and use a SoftMax output layer to assign probabilities to each category. As for age detection,

it can be quite challenging to estimate a person's exact age from their face, even for humans. Therefore, our algorithm simplifies this problem by dividing ages into eight categories, each represented by a specific range, such as (0, 2), (4, 6), and so on. These categories provide a more manageable framework for age prediction, acknowledging that factors like wrinkles, grey hair, and the "baby face" phenomenon can make precise age estimation difficult.

Deep neural network (DNN) face detector model: this section explains the technical aspects of detecting faces and predicting gender/age using the DNN face detector model. The DNN face detector model uses the single-shot-multibook detector architecture optimized for real-time object detection. Its ResNet-10 backbone architecture can capture complex features in images. For gender prediction, the algorithm processes detected faces through the gender network loaded into memory. By selecting the output node (male or female) with the highest probability, the algorithm determines the gender. Similarly, the age network predicts the age range that best fits the facial features detected for age prediction. IMSHOW from OpenCV can display the network's output on input images.

Percentage of faces discovered: this section explains how the algorithm measures bias. It does this by counting the number of faces detected within different categories and tracking the number of females, males, and faces within each age category. The counters are then used in equations to calculate the percentage of bias. For example, the female ratio is found by dividing the detected females by the total number of classified faces (excluding unidentified genders). Similar calculations determine the male ratio and the distribution of faces among age categories. This section systematically assesses the algorithm's accuracy and potential biases in predicting gender and age.

3.3. Case study on algorithm bias

The text below explores real-world examples of bias in AI systems regarding skin color, gender, and age. It highlights the significant consequences of these biases and presents case studies. For example, it discusses a healthcare algorithm that demonstrated bias against black patients, leading to unequal medical care. It also addresses gender bias in search engines, hiring algorithms, and facial recognition systems. Additionally, the section covers age bias in AI, affecting housing, employment, healthcare, finance, and banking, particularly for older individuals. The text mentions studies and campaigns from organizations such as the World Health Organization, emphasizing the importance of addressing age discrimination in AI. The overview also discusses the algorithm's approach to gender and age detection and its strategy for measuring bias. It emphasizes the importance of understanding and addressing algorithmic bias through real-world case studies, highlighting its societal impact.

4. RESULTS AND DISCUSSION

This section investigates bias in AI algorithms regarding skin color, gender, and age. Researchers developed algorithms to detect and mitigate bias, finding a bias towards lighter skin tones and young adults in the datasets tested. The findings highlight the importance of using diverse datasets to reduce bias in AI.

4.1. Implanted dataset

To perform color skin bia detection, we employed two distinct datasets:

- Dataset [24]: this dataset comprises images featuring individuals of different ages, genders, and skin tones, including black and white individuals. We obtained this dataset from Kaggle, which Jangra [24] curated. All the images in this dataset are small and colored in RGB. To make the dataset consistent, we resized all the images to 100×100 pixels, resulting in a dataset of approximately 6,000 images.
- Dataset [25]: to identify gender and age bias, we used a separate dataset. The dataset consists of actual photographs of people of different ages and genders. These high-quality images are 600×600 pixels and were obtained from the CIPLAB Institute at YONSEI University, available on Kaggle [26]. The dataset contains approximately 1,500 authentic images.

4.2. Results of the skin color bias algorithm

We utilized a skin color bias detection algorithm to analyze a dataset of 5,909 images and documented the results in Table 1. The algorithm effectively detected 5,471 faces, resulting in a highly satisfactory detection rate. Upon further examination, we discovered bias percentages within the dataset, revealing that the algorithm was biased towards white skin. Approximately 73% of the dataset featured white faces, while the remaining 27% depicted black faces. It is important to note that the constraints of the dataset may limit the accuracy of these percentages.

Nonetheless, they still offer valuable insights into bias, as depicted in Figure 2. The images are renamed based on skin color to save the algorithm's classification results, which are here. The error rate is

examined across 100 randomly chosen images, with 76 depicting individuals with brown skin and the remaining 24 featuring individuals with black skin. Table 2 provides additional details of our analysis.

Table 1. Results of the detected face

Total images	Total detected faces	Detected faces ratio (%)
5909	5471	92.587578%

Table 2. The error rate in detecting black and white faces

	Image size	The detected face	Error percentage (%)
Blackface	24	22	16%
Whiteface	76	73	4%

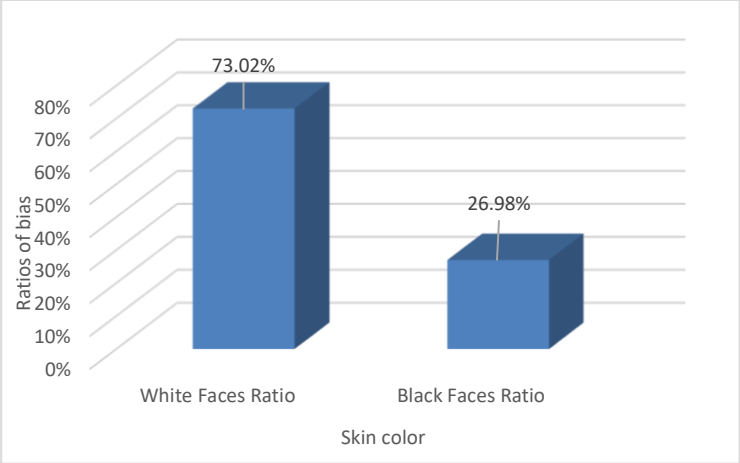

Figure 2. Ratios of bias related to skin color

4.3. Gender and age bias algorithm results

Figure 3 shows the results of running the gender detection algorithm on the dataset. The ratios indicated a nearly equal distribution between genders, suggesting no significant bias in the data. These ratios imply that the algorithm generates well-balanced gender predictions, albeit slightly favoring males. Now, let's examine the initial results on age bias ratios in Figure 4. The figure categorizes age groups, and the mentioned ratios can be used as a benchmark to evaluate age bias. However, a more comprehensive analysis is needed to confirm if there are any significant biases in the age predictions made by the algorithm. Figures 5 depict the results of applying the color, gender, and age bias ratio algorithm to the first and second datasets. They illustrate a significant improvement.

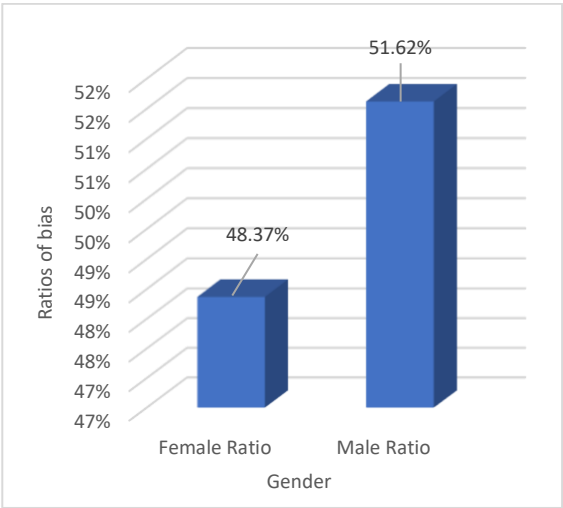

Figure 3. Preliminary results for predicting gender bias ratios

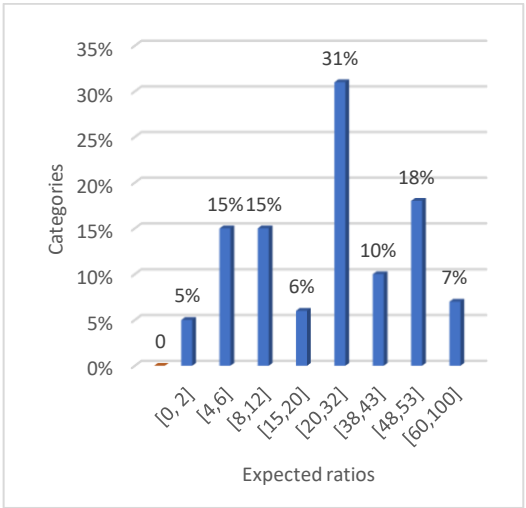

Figure 4. Preliminary results for predicting age bias ratios

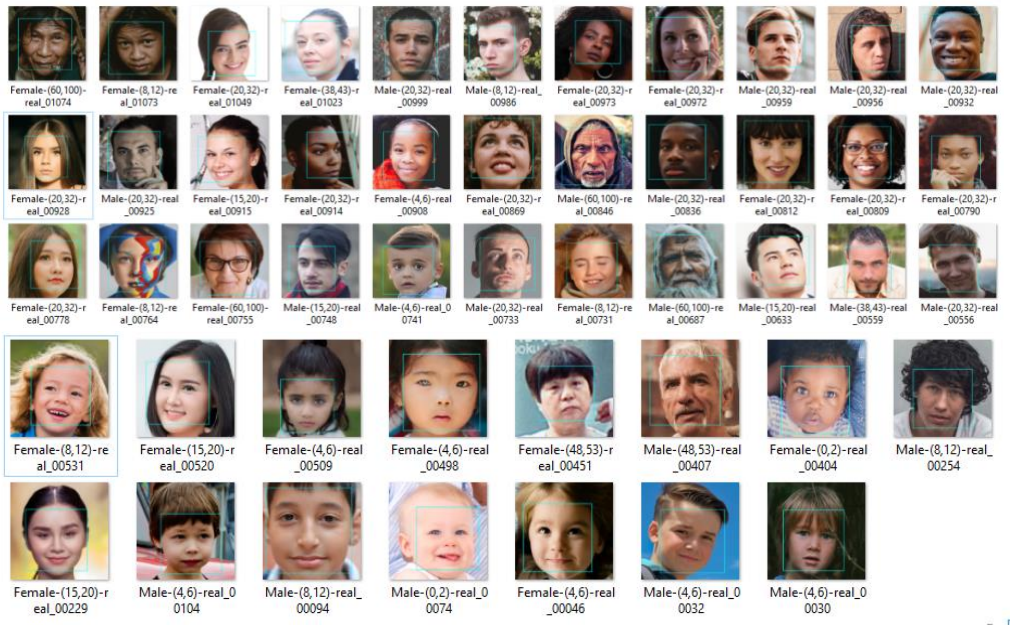

Figure 5. The result of applying the gender and age detection algorithm

4.3.1. Error rate analysis

A set of 100 random images was chosen to determine the error percentage, and the results demonstrate that 49 images had female faces and 51 featured male faces. The photos were selected from different age categories, including six images from the age group (0-2), 17 images from (4-6), 11 images from (8-12), ten images from (15-20), 39 images from (20-32), five images from (48-53), and four images from (60-100). You can find the details in Tables 3 and 4.

Table 3. The error rate in detecting female faces

	Image size	The detected female faces	Error percentage (%)
Female faces	49	46	6
Male faces	51	47	8

Table 4. The error rate in detecting age categories

Age categories	The actual age categories face	The detected age categories face	Error percentage (%)
(0-2)	6	5	17
(4-6)	17	14	18
(8-12)	11	10	10
(15-20)	10	8	20
(20-32)	39	36	8
(48-53)	5	5	0
(60-100)	4	4	0

4.4. Discussion

We developed two algorithms to detect and address bias within AI systems. The first algorithm was designed to identify bias related to skin color. We tested it on a dataset that included people of various skin tones and found it highly accurate, with a success rate of 92%. However, the algorithm also revealed a significant disparity between different skin colors. The dataset contained predominantly white faces (73%) compared to black faces (27%), indicating a bias favoring lighter skin tones. We randomly selected 100 images from the dataset to test the algorithm's performance and found an error rate of 16% for black faces and 4% for white faces.

For the second algorithm, we developed a system to detect gender and age bias. We tested this algorithm on datasets comprising images of individuals across various age groups and genders. The gender detection aspect showed a balanced distribution, with approximately 52% males and 48% females, indicating

no inherent bias. However, the age detection algorithm identified biases within specific age groups, particularly favoring young adults (20-32 years old) and exhibiting biases against children (0-2 years old) and older adults (60-100 years old). We found limited images available for children and older people, contributing to this bias.

To evaluate the accuracy of the gender and age detection algorithm, we randomly selected 100 images for testing, comprising 49 female faces and 51 male faces. The algorithm had an error rate of 6% for female faces and 8% for male faces. For age detection, we tested the algorithm on different age groups. We found that it accurately identified 5 out of 6 images in the 0-2 age group, 14 out of 17 images in the 4-6 age group, 10 out of 11 images in the 8-12 age group, 8 out of 10 images in the 15-20 age group, and 36 out of 39 images in the 20-32 age group. The algorithm demonstrated remarkable accuracy for the age groups 48-53 and 60-100, with an error rate of 0% and 4%, respectively.

5. RECOMMENDED AI ETHICAL CONSIDERATIONS AND REGULATIONS IN JORDAN

Ethical considerations and regulations are recommended for developing AI in applications to regulate the rapid development of AI in Jordan and mitigate its potential negative impacts. The recommended ethical considerations and regulations are the following,

a) Transparency and clarity: developers and organizations must be transparent about the source of data, the results obtained from it, the functioning of their algorithms, and their ultimate goals. This transparency allows for the tracking and understanding AI systems, especially regarding errors or mistakes.
b) Fairness and impartiality: data sets used for AI training must be free from bias related to skin color, gender, or race. Fairness should be considered when collecting data and conducting research. Researchers must maintain self-awareness for objectivity.
c) Prevention of misuse: AI systems must be designed with clear intentions and carefully consider potential risks if used in unintended ways as potential for security.
d) Responsibility: developers and manufacturers of AI algorithms should take responsibility for the technology they create to ensure AI systems' safe and ethical use.
e) Data privacy: data privacy is paramount for AI as it relies heavily on data for training. Therefore, responsible data collection, management, and usage must be followed. Any unnecessary data must be deleted to protect individuals' privacy.

5.1. Benefits of artificial intelligence ethics for Jordan

AI is rapidly transforming various sectors, offering numerous opportunities for innovation and efficiency. However, deploying AI technologies also brings significant ethical challenges that must be addressed to ensure their responsible use. Implementing robust AI ethics frameworks for Jordan can provide substantial benefits across multiple domains. Ethical regulations not only enhance accountability and transparency but also promote fairness, protect data privacy, and prevent the misuse of AI. These measures are crucial for fostering trust, ensuring just outcomes, and safeguarding individuals' rights, ultimately contributing to the sustainable and equitable development of AI technologies in Jordan; the benefits are as follows:

a) Accountability: ethical regulations are essential for holding AI developers accountable for their technologies and mitigating biases, errors, and unintended consequences. Transparent AI systems inspire user trust and confidence in the applications they utilize.
b) Transparency: AI systems gain user trust through transparency in their goals and decision-making processes. This openness allows users to evaluate the technology better and informs public policy development.
c) Fairness: training AI with unbiased data is crucial for ensuring fairness and making just decisions in law enforcement, hiring, and financial services.
d) Data privacy: ethical considerations and legal regulations are vital for protecting individuals' rights to data privacy and preventing companies and institutions from exploiting personal data.
e) Prevention of misuse: ethical guidelines and regulations are necessary to deter the abuse of AI in various domains, including political, cultural, and educational fields, and to prevent unethical practices such as spreading propaganda or manipulating elections.

5.2. Challenges and considerations

The proposed regulations for AI ethics come with several benefits, but there are also a few challenges and considerations to keep in mind. These include:

a) Rapid technological development: AI is rapidly evolving in various fields, and it can be challenging to keep up with these developments and adapt ethical considerations and law regulations accordingly.
b) Understanding intentions: it is often difficult to discern the true intentions behind the development and use of AI systems, making it challenging to develop ethical considerations and regulate law AI applications effectively.

c) Impact on innovation: stricter ethical considerations and law regulations may slow the pace of AI development, especially in emerging technology markets in developing countries such as Jordan.
d) Stakeholder concerns: the absence of careful ethical considerations and law regulations for AI in Jordan, including in other developing countries, highlights the importance of creating such ethical considerations and law regulations.
e) Employment decisions: AI-driven employment or termination decisions could be influenced by ethical considerations and law regulations, which might complicate human resources processes and create user concerns.

In conclusion, the ethical considerations and law regulations for AI ethics in Jordan aim to ensure transparency, fairness, and responsible use of AI technologies and applications. While they bring significant benefits, keeping up with technological advancements and understanding the true intentions behind AI development and use can be challenging. Developing AI applications requires balancing innovation with ethical considerations and legal regulations.

6. CONCLUSIONS AND FUTURE WORK

The rise of AI brings opportunities and challenges, including biases, privacy concerns, and job disruptions. Many countries have implemented ethical considerations and legal regulations to govern AI. Jordan is integrating AI into its education and industries but lacks comprehensive ethical regulations. The research aims to provide ethical considerations and legal rules for responsible AI, including algorithms to detect bias in AI systems. The goal is to create an accountable and equitable AI ecosystem. Furthermore, the researchers aim to explore the application of ethical considerations and legal regulations for AI in future internet research methodologies and digital social science research, particularly in ethnography research.

REFERENCES

[1] UNESCO, "Recommendation on the ethics of artificial intelligence - UNESCO digital library," *The United Nations Educational, Scientific and Cultural Organization*, 2021, [Online]. Available: https://unesdoc.unesco.org/ark:/48223/pf0000381137.
[2] EPRS, "The ethics of artificial intelligence: issues and initiatives," *European Parliamentary Research Service*, 2020, Accessed: Feb. 13, 2024. [Online]. Available: https://data.europa.eu/doi/10.2861/6644
[3] R. M. Kidder, "How good people make tough choices: Resolving the dilemmas of ethical living - chapter one," *Institute for Global Ethics*, pp. 12–29, 2005.
[4] "Memorandum of understanding to spread and develop the culture of artificial intelligence in Jordan," *Ministry of Digital Economy and Entrepreneurship Bayader Wadi Al-Seer*, 2024. Accessed: Feb. 13, 2024. [Online]. Available: https://www.modee.gov.jo/AR/NewsDetails/مذكرة_تفاهم_لنشر_وتطوير_ثقافة_الذكاء_الاصطناعي_في_الأردن.
[5] A. A.-Shaikh, B. A. Mahafzah, and M. Alshraideh, "Hybrid harmony search algorithm for social network contact tracing of COVID-19," *Soft Computing*, vol. 27, no. 6, pp. 3343–3365, 2023, doi: 10.1007/s00500-021-05948-2.
[6] L. Floridi and J. Cowls, "A unified framework of five principles for ai in society," *Philosophical Studies Series*, vol. 144, pp. 5–17, 2021, doi: 10.1007/978-3-030-81907-1_2.
[7] J. N. Hooker and T. W. Kim, "Toward non-intuition-based machine and artificial intelligence ethics: a deontological approach based on modal logic," *AIES 2018 - Proceedings of the 2018 AAAI/ACM Conference on AI, Ethics, and Society*, pp. 130–136, 2018, doi: 10.1145/3278721.3278753.
[8] Y. Alshamaila *et al.*, "An automatic prediction of students' performance to support the university education system: a deep learning approach," *Multimedia Tools and Applications*, vol. 83, no. 15, pp. 46369–46396, 2024, doi: 10.1007/s11042-024-18262-4.
[9] A. Jobin, M. Ienca, and E. Vayena, "The global landscape of AI ethics guidelines," *Nature Machine Intelligence*, vol. 1, no. 9, pp. 389–399, 2019, doi: 10.1038/s42256-019-0088-2.
[10] M. Kearns and A. Roth, "The ethical algorithm: the science of socially aware algorithm design," *Perspectives on Science and Christian Faith*, vol. 73, no. 1, pp. 55–56, 2021, doi: 10.56315/pscf3-21kearns.
[11] J. Kovač, P. Peer, and F. Solina, "2D versus 3D colour space face detection," *Proceedings EC-VIP-MC 2003 - 4th EURASIP Conference Focused on Video / Image Processing and Multimedia Communications*, vol. 2, pp. 449–454, 2003, doi: 10.1109/VIPMC.2003.1220504.
[12] G. Levi and T. Hassncer, "Age and gender classification using convolutional neural networks," *IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops*, vol. 2015, pp. 34–42, 2015, doi: 10.1109/CVPRW.2015.7301352.
[13] UNESCO, "Recommendation on the ethics of artificial intelligence," *The United Nations Educational, Scientific and Cultural Organization*, 2022, doi: 10.1111/bioe.12716.
[14] H. Roberts, J. Cowls, J. Morley, M. Taddeo, V. Wang, and L. Floridi, "The Chinese approach to artificial intelligence: an analysis of policy, ethics, and regulation," *AI and Society*, vol. 36, no. 1, pp. 59–77, 2021, doi: 10.1007/s00146-020-00992-2.
[15] E. Southgate, K. Blackmore, S. Pieschl, S. Grimes, J. McGuire, and K. Smithers, "Artificial intelligence and emerging technologies in schools: a research report," *Analysis & Policy Observatory (APO)*, Department of Education and Training (Australia). 2018, [Online]. Available: https://www.dese.gov.au/download/4637/ai-schools-report/6934/document/pdf.
[16] E. Southgate, "Artificial intelligence, ethics, equity and higher education: a 'beginning-of-the-discussion' paper," *National Centre for Student Equity in Higher Education*, Curtin University and University of Newcastle, 2020, [Online]. Available: https://www.ncsehe.edu.au/publications/artificial-intelligence-ethics-equity-higher-education/.
[17] "Institute for ethics in AI Oxford leading the way in AI ethics," *University of Oxford*, 2021, [Online]. Available: https://www.schwarzmancentre.ox.ac.uk/ethicsinai.
[18] Stanford's Institute for Human-Centered Artificial Intelligence (HAI), "Ethical challenges of AI applications," in *Artificial Intelligence Index Report*, pp. 126-134, 2021.

[19] T. Esther and A. Tella, "A review of the ethics of artificial intelligence and its applications in the United States," *International Journal on Cybernetics & Informatics (IJCI)*, vol. 12, no. 6, pp. 122–137, 2023.
[20] H. V. -Pekka *et al.*, "The role of explainable AI in the research field of AI ethics," *ACM Transactions on Interactive Intelligent Systems*, vol. 13, no. 4, 2023, doi: 10.1145/3599974.
[21] G. Singh, A. Mishra, C. Pattanayak, A. Priyadarshini, and R. Das, "Artificial intelligence and the institutional ethics committee: a balanced insight into pros and cons, challenges, and future directions in ethical review of clinical research," *Journal of Integrative Medicine and Research*, vol. 1, no. 4, 2023, doi: 10.4103/jimr.jimr_30_23.
[22] J. Hallamaa and T. Kalliokoski, "AI ethics as applied ethics," *Frontiers in Computer Science*, vol. 4, 2022, doi: 10.3389/fcomp.2022.776837.
[23] M. Alshraideh, A. A. J. A. -Zayed, M. Leiner, and I. M. Aldajani, "Beyond the scoreboard: a machine learning investigation of online games' influence on Jordanian university students' grades," *Applied Computational Intelligence and Soft Computing*, vol. 2024, 2024, doi: 10.1155/2024/1337725.
[24] A. Jangra, "Face mask detection ~ 12K images dataset," *Kaggle*, 2021. Accessed: Feb. 13, 2024. [Online]. Available: https://www.kaggle.com/datasets/ashishjangra27/face-mask-12k-images-dataset.
[25] EIOPA, "Artificial intelligence governance principles: towards ethical and trustworthy artificial intelligence in the European insurance sector," *European Insurance and Occupational Pensions Authority*, 2021. Accessed: Feb. 13, 2024. [Online]. Available: https://www.eiopa.europa.eu.
[26] CIPLAB, "Real and fake face detection (ed.)," *Kaggle*, 2019. Accessed: Feb. 13, 2024. [Online]. Available: https://www.kaggle.com/datasets/ciplab/real-and-fake-face-detection.

BIOGRAPHIES OF AUTHOR

Shurooq Mnawer Ibrahim received her bachelor's degree in computer science from the Hashemite University in Jordan. She obtained a master's degree in computer science from the University of Jordan in Jordan. She presented a thesis entitled ethical conducts & regulations for using AI applications in Jordan. Currently, he is seeking a Ph.D. scholarship in artificial intelligence or computer science. Her research interests include artificial intelligence, bias in artificial intelligence systems, artificial intelligence algorithms, and detection algorithms**.** She can be contacted at email: SRO8190285@ju.edu.jo or shrurooqmnawer@gmail.com.

Mohammad Aref Alshraideh currently holds the position of Professor specializing in Software Testing and Computational Intelligence at Lusail University's Information Technology College. Before joining Lusail University, he was affiliated with the Department of Artificial Intelligence at the University of Jordan in Jordan. He earned his Ph.D. in Software Engineering with a focus on testing from the University of Hull in the United Kingdom in 2007. He boasts an extensive 15-year career in the IT industry, encompassing roles such as software system design and development and project management. Throughout his professional journey, he has served in various administrative capacities, including Head Director Assistant for Computer Technology at the University Hospital, Human Resource Director, Registrar General at the University of Jordan, and Dean of the Graduate School. Moreover, he has held the distinguished Vice President for Administrative and Financial Affairs position at the University of Jordan. He has received numerous academic project grants and has made significant contributions to his field, with over fifty published papers in software testing, artificial intelligence, data mining, and digital humanity. He actively engages in workshops, seminars, and conferences within software engineering and computational intelligence. His research interests encompass software testing, artificial intelligence, data mining, and digital humanities. He can be contacted at email: mshridah@ju.edu.jo or malshraideh@lu.edu.qa.

Prof. Dr. Martin Leiner studied at Christian Theology and Philosophy in Tübingen and Heidelberg. He held positions at the University of Mainz/Germany (1990-1998), Neuchâtel/Switzerland (1998-2002), Jena/Germany (since 2002), and Stellenbosch/South Africa (as an extraordinary professor since 2017). Guest professorships led him to Kinshasa, Geneva, Fribourg, Kyoto, Washington/DC, and Prague. From 2000 to 2004, he served as director of the ethics center of the University of Geneva; since 2013, he has been the founding director of the Jena Center for Reconciliation Studies (JCRS), since 2018, the scientific director of the Academic Alliance for Reconciliation Studies in the Middle East and North Africa (AARMENA) and since 2020 the founding president of the International Association for Reconciliation Studies (IARS). In his research, he specialized in ethics and reconciliation studies. Recent publications include alternative approaches in conflict resolution; reconciliation, heritage and social inclusion in the Middle East and North Africa; and encountering the suffering of the other. He can be contacted at email: martin.leiner@uni-jena.de.

Dr. Iyad Muhsen Aldajani a distinguished researcher, holds the Executive and Research Director position at the Academic Alliance for Reconciliation in the Middle East and North Africa (AARMENA). He specializes in research applied phronesis in development for reconciliation, digital humanities, and peace studies. He coordinates the doctoral school for reconciliation, conflict transformation, and peace studies at Friedrich Schiller University and has received the Albert Schweitzer Golden Medallion for Science and Arts. He is also an author and editor of several notable books in his field and holds executive certificates from Harvard Kennedy School and Oxford Said Business School. His work focuses on data science for reconciliation and peacebuilding, and he is interested in digital humanities. He can be contacted at email: ayad.dajani@gmail.com.

Bettaz Ouarda is an Assistant Professor of Artificial Intelligence at the College of Information Technology, Lusail University. She obtained a Ph.D. in computer sciences (artificial intelligence and advanced databases) from the University of Science and Technology – Houari Boumediene in Algeria, where she has been involved as a researcher in the Laboratory of Intelligent Computing, Mathematics and Applications. She has obtained her M.Sc. in Computer Sciences from Amman Arab University in Jordan and a B.Sc. in Computer Sciences from The University of Jordan. Before arriving at Lusail University, she was a lecturer at Liverpool John Moores University for two years; before that, she was an Assistant Professor at École Normale Supérieure de Kouba in Algeria since 2007. Her research fields of interest are artificial intelligence, knowledge-based systems, description logic, access control, machine learning, and data science. She can be contacted at email: obettaz@lu.edu.qa.