IEEE Access
Multidisciplinary ⁝ Rapid Review ⁝ Open Access Journal

Received 5 May 2023, accepted 11 June 2023, date of publication 20 June 2023, date of current version 29 June 2023.

Digital Object Identifier 10.1109/ACCESS.2023.3287861

SURVEY

Unlocking the Potential of Wind Energy With Machine Learning-Based Avian Detection: A Call to Action

MARC PRINCIPATO[1], LISA HASSELWANDER[1], MICHAEL STANGNER[1], AND RICARDO BUETTNER[1,2], (Senior Member, IEEE)

[1]Chair of Information Systems and Data Science, University of Bayreuth, 95447 Bayreuth, Germany
[2]Fraunhofer Institute for Applied Information Technology (FIT), 95444 Bayreuth, Germany

Corresponding author: Ricardo Buettner (buettner@ieee.org)

This work was supported in part by the Deutsche Forschungsgemeinschaft (DFG), German Research Foundation under Grant 491183248; and in part by the Open Access Publishing Fund of the University of Bayreuth.

ABSTRACT This systematic literature review explores the potential of machine learning-based approaches to detect and prevent bird collisions with wind turbines. It provides a comprehensive review of the current approaches and identifies critical gaps in the literature, which may serve as the groundwork for future research and development in this area. As a result, this work highlights the importance of inter- and transdisciplinary cooperations.

INDEX TERMS Energy transition, environmental conservation, wind energy, machine learning.

I. INTRODUCTION

The climate change is one of the most pressing issues of our time and is a complex and multifaceted issue that has far-reaching impacts on our planet [1].

The burning of fossil fuels, deforestation, and other human activities are releasing large amounts of greenhouse gases into the atmosphere, causing global temperatures to rise and leading to a wide range of negative effects, including rising sea levels, more frequent and severe extreme weather events, and changes in precipitation patterns [2]. The Intergovernmental Panel on Climate Change in its fourth assessment report, stated that it is extremely likely that human activities, particularly the burning of fossil fuels, are the dominant cause of observed warming since the mid-20th century [3].

The impacts of climate change are already being felt around the world, with many regions experiencing more severe droughts, floods, heatwaves and storms. These impacts have significant economic and social costs, and threaten the well-being of both people and ecosystems [4]. The United Nations Framework Convention on Climate Change in the Paris Agreement, aims to limit global warming to well below 2 degrees Celsius above pre- industrial levels, and to pursue efforts to limit the temperature increase to 1.5 degrees Celsius [5].Therefore, it is important to reduce the emission of climate-damaging greenhouse gases [6].

Since the energy sector is a major emitter of carbon emissions in several countries, numerous nations have set ambitious energy transition targets in order to mitigate climate change and the impact on nature [5], [7], [8]. The expansion of renewable energies, as a replacement for conventional energy production methods, is hereby envisioned to cut emissions in the energy sector drastically [9]. Wind energy, in particular, has become increasingly popular due to its ability to generate electricity without producing greenhouse gas emissions. However it requires a significant restructuring of the power sector, including large-scale construction of new wind farms and power lines [10]. The growth of wind energy development has therefore raised concerns about its impact on wildlife, particularly birds. While wind turbines offer significant environmental benefits, the negative impact they can have on bird populations cannot be ignored because the rotating

The associate editor coordinating the review of this manuscript and approving it for publication was M. Shamim Kaiser.

blades of wind turbines can cause bird fatalities, and changes in habitat and behavior [11]. Given the rapid expansion of wind energy, this issue raises important questions about the sustainability and ethical implications of wind energy development [12].

Efforts to minimize the impact of wind turbines on bird populations have included the use of radar systems to detect bird movements [13], the modification of turbine design to make them more visible to birds [14], and the implementation of shutdown protocols during periods of high bird activity [15]. However, prolonged shutdowns of wind turbines lead to a loss of flexibility and productivity in power generation, which has been the main goal of renewable energy transition from the outset [16]. For instance, curtailing wind turbines during night to protect endangered bats can lead up to 10% loss of annual energy production [17]. Furthermore, it often not possible to allow for the expansion of windfarms into areas with high wind energy potentials as these are the main routes of traveling birds and despite these efforts, bird fatalities continue to occur [12].

Although solution approaches with human observers have been so far an efficient way to reduce mortality rates of certain species by 50% with a reduction of total energy production by only 0.07%, they are rather expensive, not available in remote areas, and less effective in bad sight conditions, e.g., during the night [18]. Consequently, "Curtailing wind turbine operation is one of the only mitigation approach proven effective at reducing wildlife mortality" [19], implicating the need for an innovative and cost-efficient way to combine ecological and zoological objectives.

In this context, machine learning (ML) algorithms have been of great interest in research for some time because of their ability to autonomously process large amounts of data quickly and efficiently [20]. Many works have already proposed and demonstrated the value of ML algorithms and artificial intelligence (AI) for managing the energy transition. For example, these algorithms are commonly used for monitoring smart grids [21], for energy production predictability [22], for managing the energy supply chain [23], and in many more applications across the energy sector [24]. Yet, these applications aim at improving existing infrastructures and do not help to minimize the trade off between expansion of renewable energies and the synthetic interventions into nature.

Yet, Fine-Grained Image Analysis is a widely studied machine learning approach that enables algorithms to detect objects and assign them to corresponding categories. This approach performantly proves its strengths in generic image recognition such as the differentiation between birds and dogs, and the detection of birds in a wildlife landscape [25]. The ability of ML algorithms to process large amounts of data quickly and accurately makes them a promising technological advancement for identifying patterns of bird activity around wind turbines and for developing real-time solutions to mitigate bird collisions. By using ML algorithms to automate the detection of bird activity and initialization of counter-measures, it may be possible to reduce the need for manual observation and human intervention, and to improve the efficiency of wind energy production while minimizing its impact on wildlife [26].

Inspired by these recent advancements in the field of image- and video-based recognition algorithms for animal identification and species classification [27], we hypothesize that these ML algorithms may provide such way to manage the conflict between the much-needed benefits of wind energy and its drawbacks due to synthetic disturbance of avifauna. Subsequently, this work aims to answer the following research question:

What is the current state of research on machine learning approaches for mitigating bird collisions with wind turbines?

The research question seeks to understand the current state of research on the use of machine learning approaches to mitigate bird collisions with wind turbines. As previously mentioned, this issue is critical as it represents a conflict between the need for renewable energy and the conservation of bird populations. Therefore, in this paper, we aim to provide a comprehensive review of the existing literature on the impact of wind turbines on birds and potential strategies to mitigate this impact through ML techniques.

To achieve this goal, we will conduct a thorough examination of the current literature on ML algorithms for detecting avifauna, focusing on common concepts and summarizing the findings. We will then discuss the challenges that arise when applying these algorithms to the energy sector. Lastly, we will identify areas for further research aimed at addressing the challenges identified in our literature review and at advancing the application of machine learning for the greater good in this sector. In doing so, we hope to contribute to the ongoing debate about the role of wind energy in promoting sustainability and the need to consider the impact on wildlife.

Furthermore, we believe that this paper offers a valuable contribution to the field of renewable energy and machine learning by providing insight into current applications of ML in a relevant area of intersection. We also highlight the importance of continued research in this area, as it is crucial to identify innovative ways to address the challenges posed by renewable energy sources and promote sustainable practices. Ultimately, we hope that this paper will serve as a valuable resource for researchers and practitioners interested in exploring the intersection of renewable energy and conservation efforts.

The remainder of this work is structured as follows: section II provides the theoretical background on the current state of machine learning and wildlife preservation in the energy sector. Section III describes our research methodology. In Section IV, we present our findings and discuss them in Section V. Section VI concludes.

II. THEORETICAL BACKGROUND

A. TRADE OFF BETWEEN BIODIVERSITY PRESERVATION AND THE ENERGY TRANSITION

The increasingly evident effects of the anthropogenic climate change [28] are intensifying the pressure on policymakers to take actions [29]. Some of the most detrimental consequences include rising global surface temperatures, melting of glaciers, rising sea levels [30], the loss of biodiversity and the destruction of ecosystems [31]. Further stress is placed on the resilience of existing habitats and ecosystems through natural resource exploitation [29], excessive land use, and human-induced fragmentation of natural habitats [28] caused by synthetic interventions in the natural landscape. Climate change was caused primarily by the excessive use and combustion of fossil fuels, which resulted in the emission of large quantities of carbon dioxide, a pollutant gas that is harmful to the climate. Historically, energy production and the implementation of fossil resources in industrial processes have had a significant impact on the magnitude of the consumption of fossil resources [30]. Consequently, the energy transition towards the dominant use of carbon-neutral renewable energies is envisioned to be a key policy tool to mitigate climate change and decrease the carbon footprint [32]. Wind energy is a decisive technology in this context [33] and has been constantly expanded in Europe since 1996. The amount of wind energy produced has more than tripled in the interval from 2010 to 2020 from 154 TWh to 513 TWh in Europe. By far the largest share of capacity within Europe is installed in Germany, followed by Spain, Great Britain and France [34]. The extent to which wind capacities are expanded depends primarily on the political and social willingness, subsidy systems and, above all, the natural wind potential of the locations. The suitability of terrestrial, mountainous and coastal regions can differ greatly, making it imperative to assess each site individually to determine whether the installation of wind power is feasible from an economic perspective [35]. Consequently, the sites for efficient implementation are limited, not easily substitutable, and therefore create conflicts with other local interests.

The expansion of renewable resources is not just an individual goal of single nations. Moreover, with the European Green Deal the European Union has set itself the goal of transforming its energy infrastructure to achieve the status of a climate-neutral continent, relying largely on wind and solar energies [36]. In addition, with the Biodiversity Strategy for 2030, the EU has established a regulatory framework for all member states on the overall objectives to be pursued in order to protect the environment from the consequences of climate change and to safeguard the functionality of our ecosystem. The first two overall objectives focused on the protection of species and habitats and the conservation and restoration of ecosystems. To achieve this, the conservation status of protected species and habitats should be improved and the performance of ecosystems enhanced through renaturation, restoration of degraded ecosystems and new green infrastructure. In order to address the goals in a reasonable approach and to be able to control the implementations, a high environmental assessment effort is required [37]. It is also fundamental to enable a strategic approach to the development of renewable resources such as wind farms resulting from the implementation of the European Green Deal and facilitating the conservation of biodiversity. Especially wind power and the protection of avifauna species are in constant tension and conflict with each other. To meet the energy demand of our society and industry, without compromising ecosystems by synthetic interventions caused by the construction of generation structures [11]. This field is aggravated by large research gaps, uncertain predictions and insufficient knowledge about complex interdependencies of consequences of the climate change on the one hand [29], [30], and by uncertain effects of wind turbines on the environment of avifauna on the other hand [38], [39].

The world bird database avibase [40] reports 916 different bird species for the European continent, 64 of these are listed as globally endangered species. The population includes numerous songbirds, water birds and birds of prey. Especially the last class, hawks, eagles and vultures, are often the focus of criticism against wind turbines. Due to low population size and slow reproduction rates, they are particularly vulnerable to habitat degradation [38]. Due to their physical structure, rotor blades are often out of the visual range of birds of prey when in flight, resulting in repeated bird strikes and causalities around wind turbines [12]. As wind turbines increase in height and rotor diameter due to technological progress, the rotor blade tips rotate at a higher pace and the risk of not being recognised by the birds at all or not being detected as a danger increases [41]. Thus, the problem of constructing wind turbines refers not only to the impact on natural habitats that are valuable for breeding, but above all to the problem of bird causalities [38].

The difficulty in detecting birds due to their elusive attributes has led to the adoption of automated techniques like machine learning to overcome traditional method limitations [42]. However, for machine learning algorithms to be effectively utilized in this context, they must have a means of interaction with the environment, which requires the provision of inputs collected from the environment in the form of data and the realization of outputs in form of actions to the environment [43].

B. APPLICATION-ORIENTED MACHINE LEARNING

Machine learning is a sub-field of artificial intelligence that focuses on developing algorithms and models that can learn from data and improve their predictions and decisions over time [44]. With the increasing availability of large amounts of data and computing power, machine learning is a rapidly growing field that has made significant advances and has become one of the most important and impactful technologies of the 21st century [45].

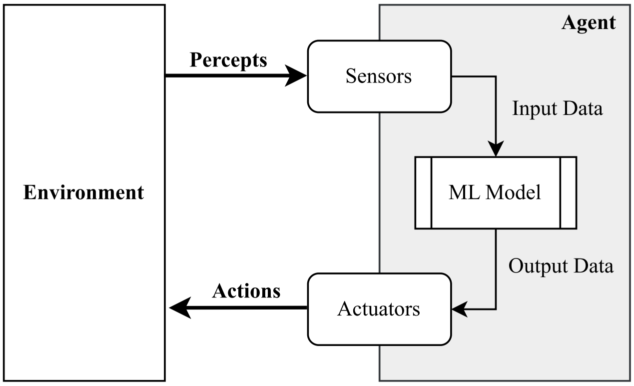

FIGURE 1. Reactive agent model in the context of machine learning.

However, one of the major challenges in this field is to develop algorithms and models that can be applied to real-world problems and deliver value to businesses, organizations, and individuals. As a framework for application-oriented machine learning, the reactive agent model (RAM) has gained growing attention [46]. This model provides a structured approach for developing and exploring machine learning algorithms specifically tailored to solve real-world problems [47]. In this context, the reactive agent model provides a framework for practical and application-oriented machine learning by facilitating the translation of machine learning research into practical solutions for various domains [44].

The RAM (see figure 1) is grounded in the concept of agents, which are autonomous entities that interact with their environment in order to achieve specific goals. Within the context of machine learning, a reactive agent can be viewed as a machine learning algorithm that interacts with its environment in real-time and makes decisions based on the information it receives [48]. The reactive agent framework describes the design and implementation of machine learning algorithms that are intended to address real-world problems [49]. In this context, the machine learning algorithm is considered as an autonomous entity, an "agent", which interacts with its environment in real time and makes decisions based on the information received from the sensors [44]. Therefore, the RAM provides a structured approach to the development and exploration of machine learning algorithms that can handle the complexity of real-world environments, while making decisions in real-time based on the information received from sensors and actions realized through actuators [50].

The agent's ability to make decisions and take actions in real-time, based on the information received from the environment through sensors, enables the algorithm to adapt to changing conditions in real-time and to make decisions in dynamic and unpredictable environments [49]. This allows the agent to develop a representation of the underlying relationships and dynamics of the domain, which can be difficult to represent through explicit rules. Additionally, the learning process can be guided by reinforcement learning algorithms, which allow the agent to learn from its own experiences and improve its behavior over time [50]. This adaptability is crucial for the success of machine learning algorithms in real-world environments, where conditions can change rapidly and unpredictably [44].

III. RESEARCH METHOD

For addressing our research question we require data that we aim to extract from published and peer-reviewed literature in this field.

Especially for exploring application potentials, systematic literature reviews (SLR) enjoy popularity in the field of ML and AI [51], [52], [53]. SLRs are a common methodological approach for collecting, summarizing, synthesizing and systematically structuring a body of knowledge in a field of inquiry [54], [55], [56], [57], [58]. Therefore, a SLR is the most suitable method approach for our research endeavour. To ensure a highly qualitative SLR, we followed the well-established SLR guidelines of Kitchenham and Charters [57] and the search strategy of Zhang et al. [59].

A. LITERATURE COLLECTION

The search strategy consists of three phases: a preliminary exploration of search string terms, the main search, and backward/forward searches. Figure 2 comprehensively summarizes our entire SLR process.

1) PRELIMINARY EXPLORATION

We conducted an initial search using GoogleScholar.com and Elicit.org for the term "machine learning" AND "recognition" AND "avifauna" to get an initial overview of the terms used in the literature on this topic. Based on the existing literature, we composed our search term of three parts: technology, purpose, and application domain (context).

- **Technology:** "machine learning" OR "deep learning" OR "artificial intelligence"
- **Purpose:** "classification" OR "detection" OR "recognition"
- **Context:** "bird" OR "bat" OR "avifauna"

Therefore, accounting for relevant synonyms, our search string is the combination of these partial strings:

(("machine learning" OR "deep learning" OR "artificial intelligence") AND ("classification" OR "detection" OR "recognition") AND ("bird" OR "bat" OR "avifauna")).

2) MAIN SEARCH AND SELECTION

As it is advised to survey multiple sources of literature to reduce bias [60], we applied this search string during the main search phase (on the 10th of December 2022) in the following most reputable databases for peer-reviewed literature on information technology, information systems, and cross-domain topics [61]:

Following the SLR guidelines [57], we defined inclusion and exclusion criteria based on filter requirements (see table 2). In the main search, we made use of the the

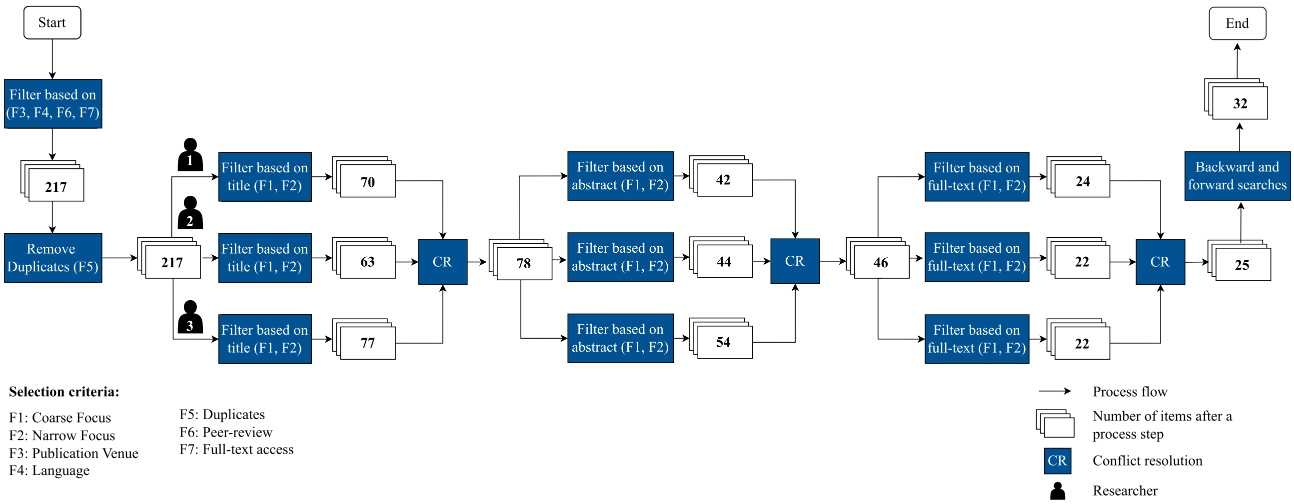

FIGURE 2. Literature selection process.

TABLE 1. Electronic data sources (EDS) used in the main search.

ID	Name	Website
EDS1	ACM Digital Library	https://dl.acm.org/
EDS2	IEEE Xplore	https://ieeexplore.ieee.org/
EDS3	Science Direct	https://www.sciencedirect.com/

TABLE 2. Filter criteria for identifying relevant items.

ID	Aspect	Inclusion criterion (AND)	Exclusion criterion (OR)
F1	Coarse focus	The item focuses on machine learning within the field of computer science AND biology	The item focuses on any other field
F2	Narrow focus	The item focuses on machine learning applications within the field of classifying flying animals	The item does not explicitly addresses this research direction OR only briefly mentions it
F3	Publication venue	Conference publication OR journal publication	The item is not of these publication types
F4	Language	The item is in English	The item is not in English
F5	Duplicates	Items are new to the search process	Item was already processed
F6	Peer-review	The item has been peer-reviewed	The item is grey literature
F7	Full-text access	The item is accessible with a University of Bayreuth account	The item is not accessible with a University of Bayreuth account

search engines' filter settings for directly conforming to our criteria of F3, F4, F6, and F7 and applied the search string to title, abstract, and keywords. The EDS search yielded a total of 217 publications across the databases. There were no duplicates to exclude (criteria F5). Subsequently, we performed three filtering steps based on criteria F1 and F2 respectively in title, abstract, and full text of each item.

For increased inter-subjectivity, the main authors performed these filtering steps independently of each other and over the entire sample (i.e., redundantly instead of splitting for increased efficiency).

Instead of having a simple binary choice of including/excluding an item, we opted to include an option for indecisiveness. This allows for a more accurate process of selecting literature by using a more detailed rating of the individual items. Concretely, this was implemented by the three researchers giving a rating of either 0, 0.5, or 1 to each item in each filtering step, whereby 0 corresponds to exclude, 0.5 to indecisive and 1 to include. The sum of votes (i.e., maximum 3 and minimum 0) of each item was then divided by 3 to obtain the mean rating. Generally, if the mean rating was above 0.5 the item was included and if it was below the threshold of 0.5 it was excluded. The list of items which had a rating of exactly 0.5 were taken into a conflict resolution (CR) round, where the authors discussed these items. Furthermore, the CR was used to discuss heavily differing votes on items and their subsequent inclusion/exclusion.

Popular approaches to calculate a measure of inter-rater reliability for multiple raters and ordinal scale are Cohen's Kappa [62] and Cronbach's Alpha [63]. However, both have been criticized in past analyses of measurements for drawbacks, such as only considering agreement, and thus not accurately reflecting the degree of reliability in cases where there is significant disagreement between raters [64] or for standardizing rater values and only measuring covariance [65]. Krippendorff's Alpha [66], on the other hand, emerges from the comparisons as a conservative, versatile and comprehensive measure [65], [66], [67].

Therefore, we make use of Krippendorff's Alpha for calculating our inter-rater reliability of the three filtering steps. In specific, we make use of the improved calculation method for a weighted Krippendorff's Alpha proposed by Gwet [68], which results in the following values:

The alpha coefficient can be interpreted similarly to other inter-rater measures, ranging from 0 (no agreement) to 1 (maximum agreement). It is customary to view an $\alpha \geq 0.8$ as good agreement between the raters, $0.8 > \alpha \geq 0.667$ as

TABLE 3. Inter-rater reliability for the filtering steps.

Filter step	K-Alpha
Title Filter	0.898030942
Abstract Filter	0.772940862
Full-text Filter	0.917865707

fair agreement, and $\alpha < 0.667$ as bad agreement [64]. Our results show a confident level of agreement among the authors in the selection of literature items, suggesting a shared understanding of the specified filter criteria.

3) BACKWARD/FORWARD SEARCHES

In the last phase, we conducted backward and forward searches, for items that might be relevant but were not yet included with our search string [58]. Newly obtained literature was then assessed against the aforementioned inclusion and exclusion criteria. We thus obtained 7 new items, resulting in 32 overall items. These thirty-two items constitute the identified and relevant literature for the body of knowledge that we aimed to review, synthesize, and structure for addressing our research question [57].

B. INFORMATION EXTRACTION

For extracting the needed information from the obtained literature items, we performed structured extraction based on filling out extraction cards [69]. The extraction comprises the following 13 fields about the literature items: Authors, year, country, publication channel, publication type, research aim, research question, research approach, covered machine learning approaches, type of detection/classification system, study findings, challenges, and opportunities for future work. To reduce bias, three authors independently and redundantly extracted this information and then participated in a conflict resolution round, where discussions were held about potential conflicts on the extracted information. The extracted information was then synthesized on the basis of a comparative analysis between items in accordance with the SLR guidelines [57].

IV. RESULTS

A. DESCRIPTIVE OVERVIEW

Figure 3 and figure 4 descriptively illustrate the final literature selection.

First works on bird identification via machine learning techniques have been published in 2011 and 2012. However, publications before 2019 remained sporadic; being two items per year at maximum. From 2019 on, literature on this topic saw a sharp increase with publications per year rising to four in 2019, six in 2020, and finally reaching its peak with 7 in 2022. There is already one item in our literature selection that is set to be published in 2023.

Most of the academic and peer-reviewed literature is published in conference proceedings (twenty items) while the rest was published in journals (twelve items). Considering the

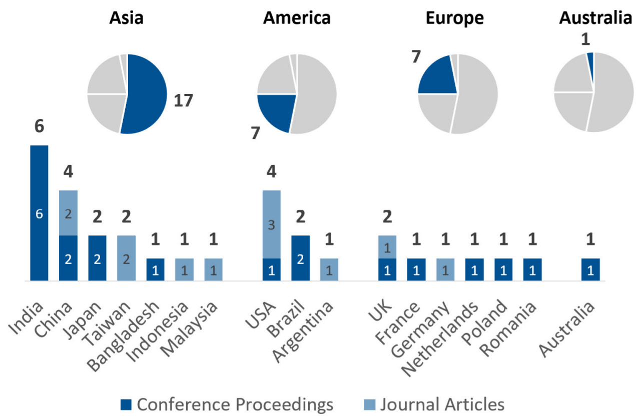

FIGURE 3. Type and origin of selected publications.

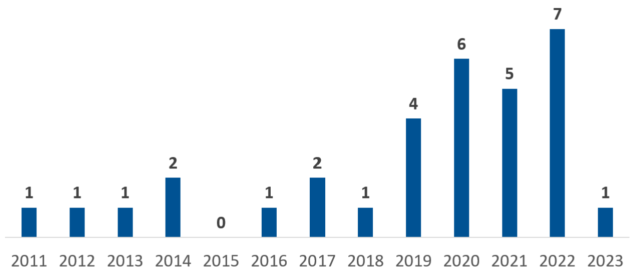

FIGURE 4. Yearly publications of bird identification approaches.

nature of academic publishing (i.e., conferences being more practice oriented with papers for concrete use-cases while journals often prefer theory-building papers and emphasize theoretical contributions), this distribution indicates that most of the research might be more practice-oriented.

The distribution of items to their country of origin seems to be dominated by Asian countries (17 items). The country of origin is hereby determined by the majority of authors and in case of tie, by the main author. The leading Asian country is India (6 items), closely followed by China (4 items). The remaining items originating from Asian countries are distributed over Japan, Taiwan, Malaysia, and Bangladesh. European countries amount to 7 publications in our final selection which are almost evenly distributed on the countries of England, Germany, France, Netherlands, Poland and Romania. American countries also amount to 7 publications in the selection which are further comprised of four items from North America (United States) and three from South America (Brazil & Argentina). Lastly, one item originated in Australia.

Concluding, annual publication numbers demonstrate that the utilization of machine learning for identifying birds and classifying bird species has gained significance as a research topic in recent years and is continuing to grow in relevance (Fig. 4). The topic enjoys considerable attention in research across the globe, however, there is a slight backlog regarding works from European and American countries.

TABLE 4. Literature items per category of the RAM framework.

Paper information			Reactive Agent Model categories					
Item	**Authors**	**Year**	**Sensors**	**Input**	**Pre-Processing**	**ML Models**	**Output**	**Actions**
[70]	Yoh et al.	2022		x	x	x	x	
[71]	LeBien et al.	2020		x	x	x	x	
[42]	Xie et al.	2023		x	x	x	x	
[72]	Chandu et al.	2020	x	x	x	x	x	x
[73]	Sharma et al.	2022		x	x	x	x	
[74]	Evangelista et al.	2014		x	x	x	x	
[75]	Incze et al.	2018		x	x	x	x	
[76]	Marini et al.	2013		x	x	x	x	
[77]	Islam et al.	2019		x	x	x	x	
[78]	Bold et al.	2019		x	x	x	x	
[79]	Kahl et al.	2021	x	x	x	x	x	
[80]	Aggarwal and Sehgal	2022		x	x	x	x	
[81]	Hidaya et al.	2021		x	x	x	x	
[82]	Yang et al.	2022	x	x	x	x	x	
[83]	Lopes et al.	2011		x	x	x	x	
[84]	Bhusal et al.	2019	x	x	x	x	x	
[43]	Nagy et al.	2020	x	x	x	x	x	x
[85]	Jadhav et al.	2020		x	x	x	x	
[86]	Boudaoud et al.	2019		x	x	x	x	
[87]	Ragib et al.	2020		x	x	x	x	x
[88]	Huang and Basanta	2021		x	x	x	x	
[89]	Kojima et al.	2016	x	x	x	x	x	
[90]	Tivarekar and Virani	2017		x	x	x	x	
[91]	Li et al.	2022		x	x	x	x	
[92]	Ruff et al.	2021		x	x	x	x	x
[93]	Stowell and Plumbey	2014		x	x	x	x	
[94]	Yang et al.	2022		x	x	x	x	
[95]	Kaminska and Gmerek	2012		x	x	x	x	
[96]	Biste et al.	2022	x	x	x	x	x	
[97]	Gupta et al.	2021		x	x	x	x	
[98]	Qiao et al.	2017		x	x	x	x	
[99]	Choudhary et al	2021		x	x	x	x	

B. CONTENT ANALYSIS

The reactive agent model can serve as an effective framework for synthesizing and organizing findings in a literature review about applied machine learning. By using this model, researchers can gain a clearer understanding of how machine learning models have been applied to solve specific problems and the conditions under which they are effective. Additionally, the reactive agent model provides a consistent approach for analyzing and comparing different machine learning models, allowing for improved understanding and the identification of similarities and differences between them. This in turn can lead to the development of more effective models and the identification of new areas of research.

To aggregate, structure and analyze our findings, we therefore make use of this framework and present our findings following the information flow in the RAM from percepts and sensors to actuators and actions.

1) PERCEPTS AND SENSORS

Following the reactive agent model, sensors play a pivotal role in the field of bird detection for this provision of inputs collected from the environment. These sensors are physical devices designed to detect and respond to physical changes or stimuli in the environment [86]. A sensor operates by converting a physical quantity, such as temperature, light, sound, pressure, or electromagnetic radiation, into a signal that then can be interpreted by a machine (i.e., a computer) [43]. In the case of ML bird identification, this means that it provides crucial data inputs to the algorithms, which then can facilitate the accurate identification and tracking of birds in their natural habitats. Furthermore, it allows for the automated collection of large data sets which is especially important for training machine learning models [42].

Sensors for bird detection and recognition mentioned in the literature can be categorized into active sensors (e.g., radars and sonars) and passive sensors (e.g., visual or acoustic sensors) [86]. The differentiation between active and passive sensors is that active sensors send out energy (e.g., radio waves for a radar) to perform measurements, while passive sensors only receive energy which they measure (e.g., sound waves for acoustic sensors). Although it is argued that active sensors are more robust against environmental conditions such as background noises [86], we do not find any approaches using active monitoring (besides mentioning them as alternatives). Passive monitoring on the other hand, is the preferred approach due to their simpler, less expensive, and more energy efficient design [42]. Passive monitoring, however, relies on bird vocalizations (acoustic sensors) or on bird images (visual sensors), which results in the sensors

TABLE 5. Type of applied sensors.

Sensor type	Sensor	Papers
Acoustic	Microphone	[43], [72], [79], [89], [96]
Visual	Camera	[82], [84]

TABLE 6. Initial input data.

Sensor type	Input data type	Number of papers
Acoustic	Audio	19
Visual	Image	9
	Video	1
Both	Image and audio	2
	Video and audio	1

needing to pick up these signals from the environment and thus including extraneous noise in the collected data. While a higher quantity and more expensive sensors can improve data quality, it often introduces a higher cost of setup due to higher sensor costs, data storage costs, and data processing costs [43], [84].

Visual sensors are often preferred because they can include a variety of information in a single image [86]. However, in circumstances where visual sensors cannot be employed due to obstruction or insufficient resolution to locate a desired object, audio-based scene analysis is a viable alternative besides using audio sensors directly [89]. The following table maps the approaches that employed sensors to their sensor types:

Most papers make use of microphone arrays for directional recordings. Chandu et al. [72] hereby recommend a microphone sampling rate of $44100Hz$ and a bit rate of either $128kbps$ or $320kbps$. Directional recordings are important for estimating the location of the bird [89], [96]. While acoustic sensors such as microphones produce sound data, visual sensors like cameras produce image or video data for the ML model [42].

Other than the aforementioned items, none consider setting up sensors for their identification approaches but rather make use of data bases with pre-recorded audio, image, and video data sets for training their models [42]. Still, the type of sensor determines the initial data type before pre-processing, even when using databases as source for input data. Table 6 illustrates the input data types and their popularity amongst the approaches.

Sensors, therefore, are a gateway from the physical to the digital world by collecting various signals from the environment and transforming them into input data for the ML agent.

2) INPUT DATA

The use of Artificial Intelligence in various application areas has increased substantially in the last decade, accelerated by advances in Machine Learning techniques. Through extensive training processes that require considerable amounts of training data, machine learning algorithms can be successfully applied to complex issues, achieving high quality predictions and precise classifications [100]. The amount of data needed to train a model sufficiently, increases exponentially with the complexity of the model. In order to be able to use the trained model outside of laboratory conditions with high accuracy rates, the training set needs to cover the full range of possible variations of features. Song and call variations of bird species may differ greatly for some individual bird species [42] and therefore will increase the demanded sample size. In the vocal analysis, it is also essential to take into account that there may be geographical variations in bird calls, and regional dialects may occur. This can lead to classification distortions [101], [102], [103], [104] in the application of the trained systems, if training and validation data sets do not originate from the same geographical region. For visual detection of birds and bats, the classification challenges vary. The image quality for detection depends largely on external parameters like the resolution due to motion blur and noise, lighting conditions [86], distance of the object and the background, as well as the environment [16], [76]. With changing weather conditions and times of the day, the individual avifauna species may appear rather heterogen [77], [86]. In addition, Boudaoud et al. [86] indicate significant variations of the flight posture and angle and differences in appearance. Disparities occur in colour or plumage, depending on the sex of the individual, the age and the season. Marini et al. [76] state that the beak is an important indicator and that the silhouette same as the posture are the most vital joint semantic components to classify birds. Once birds are fully grown, these attributes do not really change any more and are therefore reliable features for identification. The size or wingspan of many birds is sufficient to be able to assign them to species. Nevertheless, avifauna is often categorised visually according to a number of other features, such as feather plumage colour. Since the training of classification algorithms relies on existing labelled data collections [76], which might not be available or only to a small extent with regard to certain attributes.

For both detection approaches, visual and acoustic, the difficulty increases with the number of objects, classes or species to distinguish between [100]. How to differentiate between birds and non-birds, birds and other flying objects [16], and exactly how the classification process is performed: how finely the avifauna is categorized into classes, grouped by wingspan, species, subspecies, or even individuals [16], [86], [88]. With increasing complexity, the required size of the labelled data collections increases in order for the classification algorithm to achieve a high degree of accuracy.

While large data collections have emerged in many scientific fields [100], the freely accessible data sets for ornithological settings are still small in number and scope [42]. Neglecting the sources that did not specify the size of their utilized data set, 85 percent of the remaining 27 experiments had smaller sample sizes than 35,000, see Table 7. In the acoustic detection method, only datasets with

TABLE 7. Total size of the data set.

Size interval	Final identification type	Total number
<1,000	Audio, Image	6
<5,000	Audio, Image	4
<10,000	Image	4
<20,000	Image	6
<35,000	Image	3
<100,000	Image	1
>100,000	Image	2
>1,000,000	Image	1
Not specified	Image	4

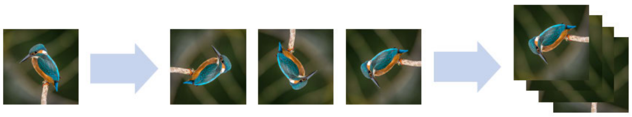

FIGURE 5. Augmentation of imbalanced data sets (image source [105]).

less samples than 5,000 were used. However, this is not justified in a lower data demand, instead it is due to the lack of large amounts of voice samples [74], [80], [83]. While larger data sets could be accessed in visual processing, see Table 7, only 3 approaches acquired data sets larger than 100,000 samples. In order to increase the data set size and to be able to achieve a better training basis for the machine learning model, it is a common approach to artificially enlarge data sets. Single images or sound recordings of a certain length are divided into smaller images or sequences. Alternatively, Huang and Basanta [88] have augmented images 10-fold by using diverse augmentation techniques such as rotating and flipping images in different directions, contrast enhancement, sharpening of images and utilizing techniques like the Gaussian noise, affine transformations or changes in the zoom range [0.7, 1.3]. Augmentation is also a suitable approach to adjust skewed data. In contrast, Aggarwal and Sehgal [80] deliberately reduced the number of their recordings to 11 percent of the original data set. Within the set was a strong imbalance between divers bird classes, leading the authors to use the minimum equal number of all classes in their approach to avoid overfitting.

Through a pre-processing procedure, data collections are partially converted into other file formats before being further processed and analyzed by the machine learning structure [73], [81]. For example, sound sequences from audio files might be transformed into Mel Spectrograms and therefore, will be further processed as an image file [43], [71], [82], [85]. The procedure and specific pre-processing methods will be discussed in the next subsection. Yet, in Table 7, the second column displays the final file format utilized. Evidently, particularly the audio data processing approaches have relied on smaller data sets. However, this might be explained by the scarcity of data [42]. Especially for species that are rare or whose habitat is mainly isolated from civilization in remote areas, it can be very difficult to obtain sufficient voice samples [70] categorized or confirmed by ornithologists.

Pretrained convolutional neural networks by leveraging existing architectures such as VGG16 have been successfully implemented in classification usecases. They present the advantages of reduced training demand, reduced computational time but still high accuracy achievements [106]. Therefore, they are also used to compensate for the deficit of training data. Table 8 shows that pretrained models are also applied in the classification of avifauna, VGG16 was utilized by Li et al. [91], Choudhary et al. [99], Gupta [97] and Islam et al. [77]. The pretrained ResNet architecture was applied by Sharma et al. [73], Li et al. [91], Huang and Basanta [88], Choudhary et al. [99], Lebien et al. [71] and Ragib et al. [87]. Li et al. [91] same as Huang and Basanta [88] adopted the Inception architecture in their models. In additon, Li et al. [91] and Chandu et al. [72] implemented AlexNet. MobileNet was deploeyed by Yang et al. [94], Choudhary et al. [99], Bhusal et al. [84]. In several approaches multiple architectures were tested and compared in the reviewed articles. Irrespective of whether pretrained architectures are used or not, large training data sets are required. These must reflect the natural diversity of attributes of the species being classified, as the performance of classification algorithms is restricted by the the tendency of the machine learning model to overfit the training data [42]. In the reviewed articles, commonly between 60 and 80 percent of the data is used for training the model and 10 to 20 percent for validation, see Table 8. In these processes, the parameters and hyperparameters are adjusted and the model is adapted and fitted [72], [88]. The amount of data used to test the model also differs between the articles. Test data (Column 6, Table 8) marked with an * are not part of the total data set (Column 3) and were reported separately by the authors. The test data set is unknown to the model or is unseen. It is used to evaluate a final model and its fit to a training data set in an unbiased approach [92].

3) PRE-PROCESSING

The main data input of the analyzed works is based on audio data. In total, 22 papers considered Passive Audio Monitoring (PAM) as a suitable method for data collection (Tab. 9), although not every approach actually followed a realistic recording approach in the wild. Similarly, with nine papers not even half as large, image input has partially initiated through realistic contexts but also via pre-annotated data sets. Only two paper considered a combined approach of collecting both, audio and image data, while one paper recorded video data which is then converted to imagery input.

However, most papers share the same pre-processing approach of collected data to gain comparable visual features. Regarding the corresponding audio data this is mainly done by converting the collected audio input into visual spectrograms, hence, using spectral data of which visualizations are easier to interpret. Thus, in total 23 papers rather consider

TABLE 8. Summary of the ML data samples by approaches.

Item	Authors	Pretrained	Total data size	Training	Validation	Testing	Species
[70]	Yoh et al.	-	34,792	70%	10%	20%	63
[71]	LeBien et al.	Y	97,900	90%	10%	1,000	24
[42]	Xie et al.	Review of Deep Learning models for acoustic bird detection					
[72]	Chandu et al.	Y	500	70%	10%	20%	4
[73]	Sharma et al.	Y	Audio: 13,700 Image: 15,180	Audio: 13,700 Image: 15,180	-	-	137
[74]	Evangelista et al.	-	1,619	90%	10%	-	76
[75]	Incze et al.	Y	-	80%	-	20%	2
[76]	Marini et al.	-	6,033	95%	5%	-	2
[77]	Islam et al.	Y	1,600	75%	25%	-	27
[78]	Bold et al.	-	Audio: 4,807 Image: 5,820	50%	-	50%	194
[79]	Kahl et al.	-	1,500,000	80%	10%	10%	984
[80]	Aggarwal and Sehgal	-	2,376	-	-	-	264
[81]	Hidayat et al.	-	752	70%	20%	10%	7
[82]	Yang et al.	-	229,164	183,690	-	45,924	264
[83]	Lopes et al.	-	312	-	-	-	3
[84]	Bhusal et al.	Y	>5,000	60%	20%	20%	-
[43]	Nagy et al.	-	-	60%	20%	20%	-
[85]	Jadhav et al.	-	700	90%	-	10%	7
[86]	Boudaoud et al.	-	12,352	80%	20%	768*	bird/not bird
[87]	Ragib et al.	Y	14,000	10,000	2,000	2,000	200
[88]	Huang and Basanta	Y	31,320	80%	20%	760*	29
[89]	Kojima et al.	-	140	-	-	-	8 categories
[90]	Tivarekar and Virani	-	200	70%	-	30%	4
[91]	Li et al.	Y	System: 18,000 Model: 600	80% 80%	20% -	- 20%	300 10
[92]	Ruff et al.	-	173,964	80%	20%	131,767*	Species: 14 Classes: 17
[93]	Stowell and Plumbey	-	10,699	-	-	-	-
[94]	Yang et al.	Y	8,057	80%	20%	-	83
[95]	Kaminska and Gmerek	-	-	70%	-	30%	10
[96]	Bistel et al.	-	3,500	2,000	1,200	100	1 (52 songs)
[97]	Gupta et al.	Y	15,032	80%	10%	10%	100
[98]	Qiao et al.	-	-	80%	-	20%	200
[99]	Choudhary et al.	Y	29,506	27,506	1,000	1,000	200

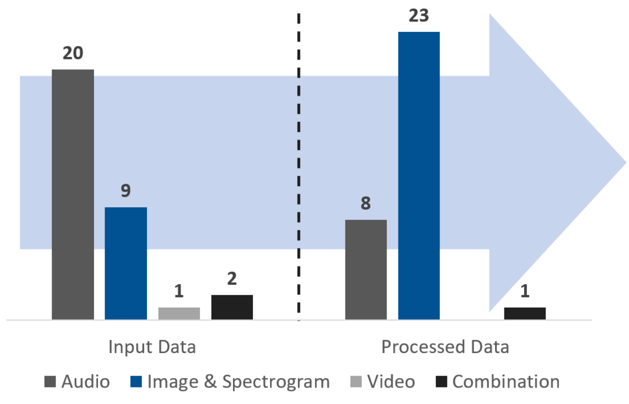

FIGURE 6. Data input source and further (pre-)processing.

TABLE 9. Transformation of input data for further (pre-)processing.

Input Data	(Pre-) Processing	Number of papers
Audio	Spectrogram	11
Image	Image	9
Audio	Audio	8
Image, Audio	Image	1
Video, Audio	Image	1
Video	Image	1
Audio	Audio, Image	1

visual elements for bird identification which in fact means that only eight of 22 papers considered a further audio processing of already collected audio data (or combined data), which is mostly done with a further processing of the mel spectrum to corresponding Mel-Frequency Cepstral Coeffcients (MFCC) (Fig. 6). Both pre-emphases enable the researchers for the amplitude calibration of audio signals according to different frequency bands and thus to standardize and prepare the data for further processing.

A common issue of audio recordings in the field is the differentiation of requested bird sounds with environmental signals or overlapping sounds of other animals and birds. Some analyzed papers only used high quality input data with a high Signal-to-Noise ratio (SNR) [73], [79], yet, this is usually not the case when relying solely on field data. Hence, there is the need to first standardize the collected data to equal file parameters (e.g. WAV format), and subsequently discriminate noise from bird syllables. This is commonly handled with a de-noising process to reduce background information to a minimum (Fig. 7). In detail, there are approaches that focus on the filtering of given audio

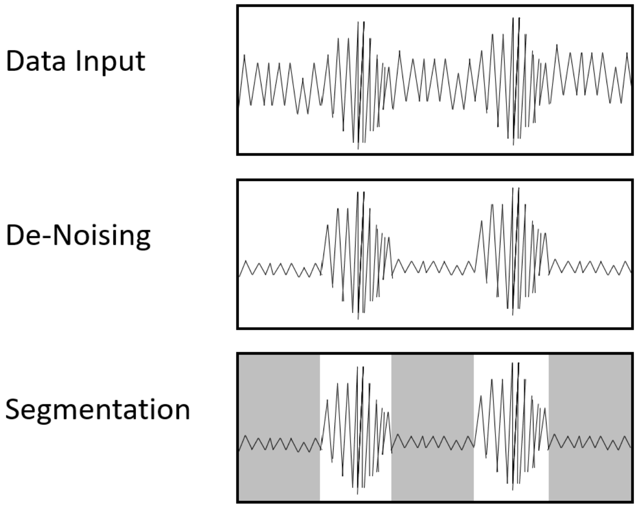

FIGURE 7. Processing of raw audio data.

signals, e.g. by setting frequency thresholds that exclude usual background noises [70], [74], [75], or by determining frequency ranges or other specific features that approximate to corresponding bird features [93], [96]. Another de-noising approach is the use of a spectral representation of the environmental noise, which is then subtracted from the audio signal to get an adjusted and de-noised presentation of corresponding bird syllables [71], [81]. Both approaches then further allow for the segmentation of relevant amplitudes by extracting these bird syllables from irrelevant sound elements that have the potential to positively influence misinterpretations (Fig. 7).

In a following step, the audio signals are commonly converted to spectrograms (Fig. 10), usually with the use of Fourier Transformations (Fig. 8). The most common algorithm within the analyzed papers is the Fast-Fourier Transformation (FFT), but also other variations such as Short-Time Fourier Transformation (STFT), Discrete-Fourier Transformation (DFT), and Discrete-Time Fourier Transformation (DTFT) have been used (Tab. 10). All have in common the transformation of large amplitude and and time dimensions of wave forms to its frequency components in Hertz (Hz) across time. While wave-visualizations depict the development of amplitudes (Y-axis) over time (X-axis), Fourier Transformations are capable of expressing the total magnitude (Y-axis) across the frequencies (X-axis) (Fig. 8). The higher the magnitude, the higher the relevance of this tone being a relevant pattern in bird detection.

In contrast to wave forms, frequency spectrograms then express the frequency at specific time points instead of the amplitude. Displayed colors, which in fact are spectral expressions of the audio signals, visualize with the level of brightness the availability and intensity of specific frequencies (Fig. 10). The advantage of these representations is the superior ability to detect and analyze varying patterns such as timbre and pitch of the bird syllables and thus its

TABLE 10. Used Fourier Transformations.

Item	Authors	FFT	STFT	DFT	DTFT
[74]	Evangelista et al.	x	-	-	-
[97]	Gupta et al.	x	x	-	-
[81]	Hidayat et al.	-	x	x	x
[75]	Incze et al.	-	x	-	-
[85]	Jadhav et al.	-	-	x	-
[79]	Kahl et al.	x	-	-	-
[95]	Kaminska and Gmerek	x	-	-	-
[71]	LeBien et al.	x	-	-	-
[73]	Sharma et al.	-	-	x	-
[93]	Stowell and Plumbey	-	x	-	-
[82]	Yang et al.	x	-	-	-
[70]	Yoh et al.	x	-	-	-

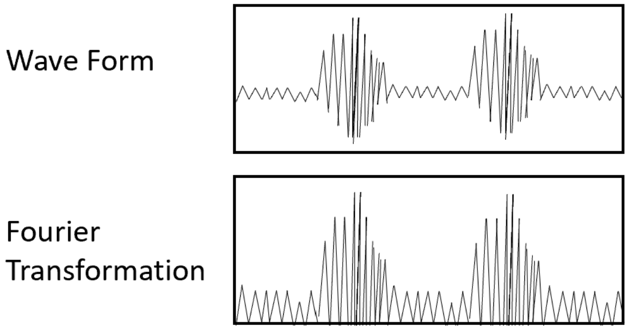

FIGURE 8. From Wave-forms to Fourier Transformations.

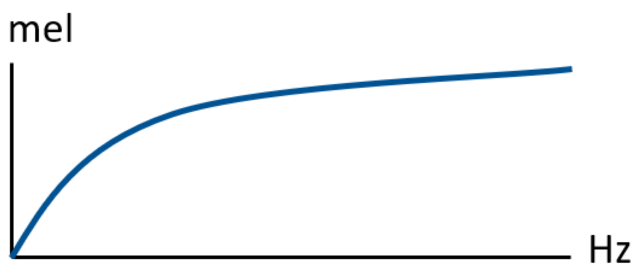

FIGURE 9. Mel Scale.

suitability for the classification of birds with an enriched visualization of the sound's structure.

A following procedure in our reviewed literature is then the use of Mel-Spectrograms (Tab. 11), which is the depiction of the time opposed to mel, instead of the frequency. Mel (m) is characterized as a logarithmic transformation of the frequency (f) in Hz:

$$m = 2595 \cdot \log_{10}(1 + f/700) \iff m = 1127 \cdot \ln(1 + f/700) \tag{1}$$

The mel scale expresses the human perception of tone, which deviates from linear spectrograms as seen before. The higher the frequency in Hz, the worse the human capability to acknowledge this difference in frequency (Fig. 9).

Thus, the difference between a linear spectrogram and a mel spectrogram is the logarithmic transformation of the linear frequency, making the latter better suitable for comprehending and analyzing changes in frequency,

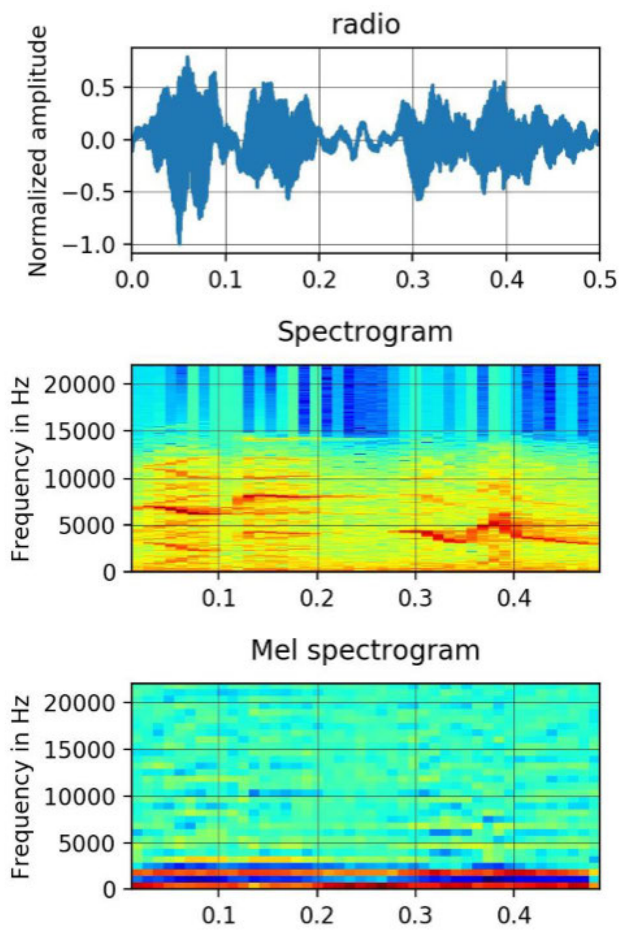

FIGURE 10. Processing of audio data to Mel-Spectrograms (source [42]).

TABLE 11. Described Mel Approaches.

Authors	Item	Mel-spectrogram	MFCC
[80]	Aggarwal and Sehgal	-	x
[74]	Evangelista et al.	-	x
[97]	Gupta et al.	x	-
[81]	Hidayat et al.	x	x
[85]	Jadhav et al.	-	x
[95]	Kaminska and Gmerek	-	x
[71]	LeBien et al.	x	-
[83]	Lopes et al.	x	x
[43]	Nagy et al.	x	-
[73]	Sharma et al.	x	-
[93]	Stowell and Plumbey	-	x
[90]	Tivarekar and Virani	-	x
[82]	Yang et al.	x	-

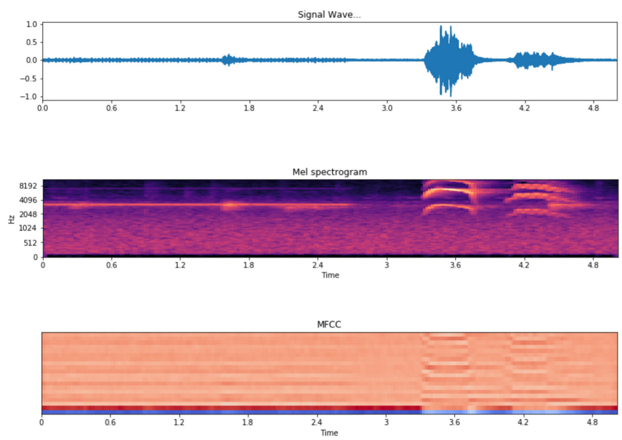

FIGURE 11. From Mel-spectrograms to MFCC (source [81]).

due to the integrated colors visualizing the intensity of tone depending on the human's ability of acknowledging frequencies (Fig. 10), which explains its common use within our analyzed papers.

A further Discrete Cosine Transform (DCT) is then applied to obtain the Mel-Frequency Cepstrum, which is built of Mel-Frequency Cepstral Coeffcients (MFCC). This approach has been been a standard in speech and audio processing for some time [107], and is also the dominant described procedure in our analyzed literature (Tab. 11). In contrast to a mel spectrogram, the mel cepstrum is a compromised and less complex depiction of the audio signals, which allows for a streamlined analysis of underlying vectorized numeric data. This enables to finally detect distinguishable sound patterns out of the feature vectors. The corresponding visualization shows accordingly a reduced heatmap containing only the audio's main components such as pitch, timbre and sound energy (Fig. 11). De-Noising and Segmentation procedures can be also applied to (Mel-) spectrograms and MFCC.

On the other hand, image data is usually pre-processed by scaling the images to identical dimensions, hence ensuring a better comparability and thus a better performance of the learning processes. Moreover, visual features have been selected on basis of color specifications (e.g. RGB, HSV, YUV) and/or on dimensions of semantic bird elements, such as head, body, tail, beak and eye, as well as the distances between these elements. Similar to the processing of audio data, deviating color segments can be extracted from the image background to analyze distinguishable visual features. This is done by removing small stripes from the border of the images, in case the bird is captured in the center of the picture [76]. These outer stripes, which make roughly 2 to 10 per cent of the picture, are being analyzed with color histograms for each base color (e.g. RGB: Red, Green, and Blue). Also other specifications such as density and saturation can be taken into account. This allows for detecting the color intensity within each pixel. The counting of pixels with specific color intensities then enables to describe the color distribution of the total image. Hence, the cropped outer stripes with their specific color distributions can be labelled as background, which allows to detect other background parts with similar pixel intensities and color distributions in the remaining image. Thus, other distinguishable parts can be labelled as bird, which then can be cropped from the unnecessary background and used for feature extraction (Fig. 12).

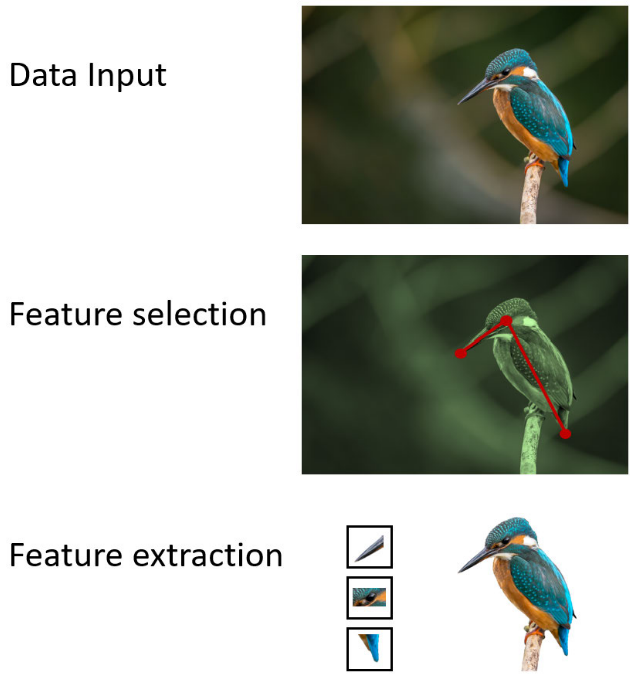

FIGURE 12. Processing of image data (image source [105]).

The features themselves can be based again, either on color distributions, on distinguishable parts of the bird, or a combination of both (Fig. 12). The color histograms of the bird itself can serve as feature vectors in a binning process that categorizes specific color distributions, hence to detect whether the bird contains rather red, green or blue pixels. Pre-annotations of these categorizations than help to classify birds with similar colors accordingly [76], [98]. Another feature extraction method is the selection of key parts of the bird. This can be done as a solitary feature extraction process [91] or in combination with a color detection process, for instance to approximate the correct classification with the use of diverse features in a decision tree [98]. Especially head, beacon, eye, trunk and tail are often claimed as decisive parts of birds that are suitable in a classification process. Dimensions of these parts, such as length, height and weight, are considered as unique elements of specific birds species and thus taken into account when categorizing detected birds. The distances between these elements are further detected to draw conclusions on body shape and total dimensions of the bird, enabling to classify the correct species.

4) MACHINE LEARNING MODELS

The data that has been pre-processed needs to be fed into a machine learning model in order to train classifiers that allow for the later identification of birds and other flying animals of the avifauna. Table 12 summarizes the analyzed approaches and their selected machine learning models for training these detection classifiers.

TABLE 12. Used Machine Learning models.

Item	CNN	SVM	kNN	RF	Other	SUM
[70]	-	-	-	x	1	2
[71]	x	-	-	-	-	1
[42]	Review of Deep Learning models for acoustic bird detection					
[72]	x	-	-	-	-	1
[73]	x	-	-	-	1	2
[74]	-	x	-	-	-	1
[75]	x	-	-	-	-	1
[76]	-	x	-	-	-	1
[77]	x	x	x	x	-	4
[78]	x	-	-	-	-	1
[79]	x	-	-	-	-	1
[80]	x	-	-	-	-	1
[81]	x	-	-	-	-	1
[82]	x	-	-	-	-	1
[83]	x	x	x	-	2	5
[84]	x	-	-	-	-	1
[43]	x	x	-	-	-	2
[85]	x	x	x	x	3	7
[86]	x	-	-	-	-	1
[87]	x	-	-	-	-	1
[88]	x	-	-	-	-	1
[89]	-	-	-	-	1	1
[90]	-	-	-	-	1	1
[91]	x	-	-	-	-	1
[92]	x	-	-	-	-	1
[93]	-	-	-	-	-	1
[94]	x	-	-	-	-	1
[95]	-	-	x	-	1	2
[96]	x	-	-	-	-	1
[97]	x	-	-	-	1	2
[98]	-	x	-	-	-	1
[99]	x	-	-	-	-	1
SUM	23	7	4	3	12	-

It becomes apparent that the most popular model by far is the Convolutional Neural Network (CNN) with 23 applications out of thirty-two possible approaches. It is followed by the traditional Support Vector Machine (SVM) which only is used by 7 approaches, the k-nearest Neighbour (kNN) algorithm with 4 uses, and the Random Forest (RF) approach with 3 overall uses. Although, there are approaches that also use other machine learning models, these methods are not often used overall (two applications at most). Further, it seems that the majority of the literature only considers one machine learning model for their solution. However, some papers survey multiple approaches such as [85] and [83].

In the following we will elaborate based on the most common four models, how the analyzed works deal with bird detection classifier training.

a: CONVOLUTIONAL NEURAL NETWORKS

A convolutional neural network (CNN) is a type of deep learning neural network that is commonly used for image classification tasks because of its specialization in processing data that has a grid-like structure [42]. The basic building block of a CNN is the convolutional layer. The convolutional layers are made up of a number of neurons, which are connected to small regions of the input image in a way that allows them to detect specific features. Each convolutional layer contains a number of kernels, which are used to scan the image and detect features. These filters are used to extract features from the image, such as edges, textures, and patterns. The output of the convolutional layer is then passed through a non-linear activation function (e.g., a ReLU), which introduces non-linearity to the network and allows it to learn more complex features [77]. The output of the activation function is then typically passed through a pooling layer, which reduces the spatial dimensions of the feature maps (i.e., reducing the image resolution), making the network more computationally efficient and allowing the network to identify more abstract features in the data, rather than just the individual pixels [85]. Once the convolutional and pooling layers have been applied, the resulting feature maps are then passed through fully connected layers, which use the extracted features to make a prediction about the class of the input image. In the case of bird classification, the CNN is typically trained on a dataset of images of birds or spectrograms of their vocalizations, with each image being labeled with the specific species of the bird that is depicted [82]. During training, the CNN thus learns to recognize the features that are unique to each species of bird, such as the shape and color of the beak or unique peaks in frequencies of their calls, and so on [97]. Once the CNN has been trained on its training dataset, it can then be used in the field to classify birds, by comparing the features that it extracts from the field data to the features which it has learned during training [84].

b: SUPPORT VECTOR MACHINES

Support Vector Machines (SVMs) are a type of supervised learning algorithm that can be used for classification and regression tasks. It is an algorithm designed to find the maximum margin hyperplane which best separates classes of data [77]. Hereby, the hyperplane is a decision boundary that separates the different classes of data by finding the maximum margin (i.e., distance) between the closest data points from either class. During the training phase, the SVM algorithm finds the hyperplane that maximally separates the data points of different classes (i.e., the maximum margin hyperplane) by maximizing the margin between the closest data points of different classes, known as support vectors [76].

In the case of bird classification, the SVM algorithm of the analyzed approaches is usually trained on a dataset of images of birds or their vocalization spectrograms, being labeled for the species that is represented by it. During the training phase, the SVM algorithm finds the hyperplane that best separates the different classes of birds based on their features [42]. Analyzed approaches hereby use the so called "kernel method": The algorithm transforms the input data into a higher-dimensional space where a linear decision boundary can be found [77]. This transformation is done by using kernel functions that help classifying data that is not linearly separable by taking the input data and mapping it into a higher dimensional space where it becomes linearly separable [42]. The kernel method therefore allows the SVM to separate the data even if it can't be separated by a straight line or a plane in the original input space. Commonly used kernel functions by our approaches are linear, polynomial and radial basis function [42], [76], [77], [85].

After finding the optimal decision boundary in training, this boundary can then be used to classify new data points in the categories (usually for a binary classification task) [42]. In the bird detection task, the trained decision boundary hereby is used to classify images of birds or their vocalizations that were obtained in field recordings (and not used for training) by determining which side of the hyperplane the new data falls on.

c: K-NEAREST NEIGHBOUR

The K-Nearest Neighbors (kNN) machine learning algorithm is a type of supervised learning algorithm that can similarly be of use for classification and regression tasks. The basic idea behind kNN is that an object can be classified by a "majority vote" of its k nearest neighbors, where k is a pre-specified parameter by the user which set up the model [77]. The algorithm works by storing all the training data, and when a new object needs to be classified, the algorithm finds the k training objects that are closest to the new object, and the new object is assigned the class that is most common among those k nearest neighbors [83].

Training a kNN algorithm is done by storing the training data and its corresponding labels (i.e., classes like the different bird species) in the memory [95]. This means that in comparison to the previously mentioned models, the kNN algorithm does not learn a model; Instead, it simply stores the training examples in memory, along with their features and corresponding labels. The algorithm learns the features of each class by analyzing the characteristics of the training examples that belong to that class (e.g., the vocalization features of bird species or their colour and beak) [95].

When new birds (i.e., field data) need to be classified, the kNN algorithm then compares the features of the new data to the features of the data in the training set. Subsequently, the algorithm then finds the k training data entries that are closest to the new image based on a distance metric and the new entry is then assigned the class that is most common among those k nearest neighbors [42]. In our analyzed approaches common metrics for the distance were the Euclidean distance or Manhattan distance [77], [95].

It is worth emphasizing again, that since a kNN is a simple, memory-based algorithm, it does not make any assumptions about the underlying probability distribution of the data (i.e., builds a model). In essence, this means that it does not make any assumptions about the functional form of the relationship between the input and output variable as other machine learning algorithms tend to do (e.g., neural networks).

d: RANDOM FOREST

Random Forests are ensemble machine learning algorithms that can, as the previous three algorithms, be used for both classification and regression tasks. Ensemble machine learning algorithms work by combining the predictions of so called "base models" (in the case of Random Forests:

multiple decision trees) to improve the accuracy and stability of the overall model [70].

In the Random Forest model, a decision tree algorithm is used to split the data into smaller subsets based on a set of rules. Each split is based on a specific feature of the data, and the goal is to create a tree structure that separates the data into different classes as accurately as possible [93]. However, decision trees have the tendency to overfit the data, meaning that they can perform well on the training data but perform poorly on new, unseen data.

Random Forest address the problem of over-fitting in individual Decision Tree classifiers by creating multiple decision trees and combining their predictions, whereby each tree is created using a different subset of the training data (chosen at random from the superset of total data) [93]. The final prediction is made by averaging the predictions of all the trees, which reduces the variance of the model and improves the stability and generalization of the model.

In the analyzed approaches Random Forest algorithms are trained on a labeled dataset of bird data, with each labeled corresponding to the specific species of the bird which is represented in the dataset [85]. As previously explained, the Random Forest algorithm creates multiple decision trees during the training phase. Each decision tree hereby using a different subset of the training data superset.

The algorithm learns the features of each species of bird and once the Random Forest algorithm has been trained, its splitting rules can then be used as classifier to categorize new bird data by averaging the predictions of all the decision trees [42].

5) OUTPUT DATA

To evaluate the performance of different recognition methods, it is necessary to define uniform evaluation metrics. The most commonly used evaluation metrics were assembled by Xie et al. [42] and are presented in this section.

The prediction results from a binary classification model for a specific class or species can be divided into four categories: True Positive (TP), True Negative (TN), False Positive (FP) and False Negative (FN). These categories are the numerical basis for the calculation of several evaluation metrics. A TP result means, that the classification of the species or class was performed correctly. A TN case identified an other class as an other class accurately. In FP or FN cases the classified object and prediction do not align. A falsely classified other species for the species indicates a FP event. In a FN case the species is mistakenly categorized for another one [42].

The accuracy of a model is the ratio of TP results and the total sample size, therefore the accuracy refers to the correctly identified cases ([TP+TN]/[TP+TN+FN+FP) [42]. The precision incorporates all predictions which are correct (TP/[TP+FP]). The ratio of the number of samples correctly predicted by the algorithm as the species to the sum of samples of the species [71]. The metric evaluates the models ability to correctly predict a class [42]. How accurate a model can categorise a class is described by the recall measurement (TP/[TP+FN]). It is calculated as the ratio of samples correctly predicted as a class to all samples recognized as this class [88]. However, the precision and recall are mutually exclusive, meaning that when one value increases, the complementary value decreases. Therefore the Fn Score (2*[[Recall*Precision]/[Recall+Precision]]) merges the summed average of precision and recall to evaluate the performance of the model conjointly [88]. A larger Fn Score implies greater performance, whereas n indicates the weighted ratio of recall and precision, calculating the bias towards the recall value [42].

In multi-class classification, common metrics us the average or sum of calculated values. The average precision is adapted from the prediction-recall-curve-integral, measuring the weighted sum of precisions. While the mean average pecision (MAP) averages the recognition accuracy of all target classes or species divided by the number of them, to evaluate the overall accuracy of a classification model [71]. Another performance evaluation measurement is the integral of the Receiver operating characteristic graph (AUC). When the horizontal axsis is defined by the FP rate and the vertical axis by the TP rate of the model, a value closer to 1 or 100% depicts a better recognition performance. Values below 0.5 or respectively 50% represent no classification abilities [42].

The most widely reported measurement in the reviewed articles is the overall accuracy of the model or for specific classified species. Table 13 shows the classes or number of species (Column 2) of the reviewed article, as well as the achieved accuracy (Column 3), precision (Column4) and other performance metrics of the different approaches, if provided by the authors.

Table 13 demonstrates varying performances among the reviewed models. In some cases, very robust models were developed which achieved a high level of accuracy even with a large number of different species or classes to be distinguished of, see Sharma et al. [73], Li et al. [91], Yang et al. [82] and Choudhary et al. [99]. Strikingly, the classification accuracy of a model can vary greatly between different avifauna species to be categorised. Particularly high precision rates were achieved by the models from Hidayat et al. [81], Huang and Basanta [88], LeBien et al. [71], Tivarekar and Virani [90].

Different evaluation metrics, suitable for the individual recognition scenarios of avifauna, depending on the chosen machine learning algorithm and the complexity of the input, are inevitable [42]. However, the absence of result standards and performance evaluation makes it difficult to compare different recognition machine learning approaches.

6) ACTUATORS AND ACTIONS

Most papers position their work by emphasizing the importance of avian bio diversity. Accordingly, several papers justify their research goal as the need for a solution for avian surveillance, and, respectively, opportunities to count and track bird populations that are threatened by extinction.

TABLE 13. Performance metrics.

Item	Species	Accuracy	Precision	Other
[70]	63	1st Model: 90% 2nd Model: 80% 3rd Model: 80%	-	-
[71]	24	-	∅ precision per class: 100-28% Total average precision: 97.5%	MAP: 89.3%
[42]	Review of Deep Learning models for acoustic bird detection			
[72]	4	90%	-	-
[73]	137	Audio Model: 92.4% Image Model: 97.1% Combined Video: 90%	-	-
[74]	76	59.75%	-	-
[75]	2	83%	-	-
[76]	2	91.18%	-	-
[77]	27	SVM: 89% kNN: 85% Random Forest: 87%	-	-
[78]	194	78.15%	-	-
[79]	984	78%	-	MAP: 79.1% AUC: 0.974
[80]	264	92%	-	-
[81]	7	Baseline Net + mel-spectogram input: 91.18% Baseline Net + MFCC input: 94.12% Dual-Input CNN + both features: 97.06%	94.33% 91.67% 95.24%	Recall: 85.71% MAP: 94.36% Recall: 89.29% MAP: 96.08% Recall: 96.43% MAP: 97.55%
[82]	264	95.12%	-	-
[83]	3	99.7%	-	-
[84]	-	Without super resolution: 70% With super resolution: >90%	-	-
[43]	-	Phase I: 84% Phase II: 96%	-	-
[85]	7	Classification of main classes: 96.7% Sub classification (based on song/call): 96-99%	-	-
[86]	bird/ not bird	95.31%	-	-
[87]	200	Basemodel: 32.78% ResNet18: 82.00% ResNet34: 83.43% ResNet50: 84.36% ResNet101: 86.43%	-	-
[88]	29	Model 1 (no swapping): 97.11% Model 2 (Swapping): 98.42% Model 1+2 (combined): 98.39% Detection of birds: 100%	98.49%	Recall: 97.50% F1-Score: 97.90%
[89]	8 categories	>60%	-	-
[90]	4	Asian Koel: 96.3% Brahminy Kite: 92.31% Grey Francolin: 100% Hill Patridge: 100%	Asian Koel: 100% Brahminy Kite: 100% Grey Francolin: 95.65% Hill Patridge: 92.86%	-
[91]	300	95%	-	-
[92]	14	99.69%	-	-
[93]	-	-	-	AUC (feature learning): 85.4% AUC (mel spectra): 82.2% AUC (MFCC): 69.3%
[94]	83	88.07%	89%	-
[95]	10	kNN: ∅69.94% SOM: ∅52.92%	-	-
[96]	1 (12 songs)	-	∅81%	Recall: 72% F1-Score: 73%
[97]	100	GRU: 67% GRU+LMU: 66%	-	-
[98]	15	83.87%	-	-
[99]	200	VGG16: 96% ResNet50: 85% MobileNetV2: 95%	-	-

With this goal in mind, there is great agreement as regards the benefits for ornithologists, however, although possible mobile or other front-end solutions are often claimed as future opportunities, there are only few papers that propose a manifest solution.

A basic solution was presented by Ruff et al. [92], who created an R-based graphical interface that allows biologists and managers to independently generate spectrograms from audio data and thus classify the species of interest in combination with the locality of the recording. This system is able to process 100h of audio data within one hour, enabling biologists via portable solutions or desktop applications to track birds with usual-performing personal computers and minimal delays.

Similarly, Chandu et al. [72] have designed a graphical user interface (GUI) to operate in real-time the underlying processes from recording in real environments over ML procedures and data processing to the displaying of results, especially for ornithology purposes. Further mobile solutions as well as their usage combination in ecological park and bird sanctuaries have been proposed as future work.

Ragib et al. [87] also have developed a web-based API service website that enables users to upload bird images. Hence, the underlying deep leaning model predicts and classifies the bird species, which is subsequently presented to the user again. Yet, the purpose of this web deployment was still focused on testing the robustness and compatibility of underlying model and not to serve as an action-oriented and problem-solving tool.

In contrast, Nagy et al. [43] have implemented a Power Bi dasbhoard that enables the user to finetune parameters, monitor the system and analyze or download resulting data. In addition, this approach includes the possibility of automatic responses via SMS or e-mail. Thus, this is the only paper that has shown the capability of an automatic reaction based on an event that has been determined as significant in advance.

Next to general bio diversity purposes, only few papers meet the goal of this review to present the tension between the energy transition and the protection of avifauna. And if they do, they barely go into detail. For instance, Boudaoud et al. [86] have claimed the need to protect marine birds due to the increased amount of wind energy plants. Others have also mentioned similar situations in which birds are endangered, such as in the vicinity of airports [74], [90]. However, although there are approaches to detect and measure bird populations, none of these provide an action-oriented, specific solution that contains an approach to avoid corresponding damages.

V. DISCUSSION

A. USING MACHINE LEARNING FOR BIRD DETECTION IN WINDFARMS

Birds are an important but vulnerable component of ecosystems and are subject to protection underpinned by European conservation law [37]. Due to the increased development of renewable energies, specifically wind energy in many regions, the habitats of birds are increasingly influenced by the synthetic generation structure. While many causal connections are still uncertain, it is important to leverage the existing technological possibilities to minimise the impact on the environment and the causalities of avifauna. As a result, the energy transition can be accelerated if the approval of wind turbines can be facilitated in the future through accurate protection options for birds and bats. This makes it feasible to interlink climate mitigation strategies with biodiversity conservation [16], [38], [41].

However, the issue of balancing the protection of avifauna with the expansion of wind power in high wind potential areas remains unresolved due to the lack of reliable and automatic methods to prevent birds from flying into the dangerous area of the rotor blades, respectively collusions of avifauna and wind turbines [16], [41]. The use of a reactive agent in the form of a machine learning approach that performs bird detection, recognition and initiation of countermeasures (if necessary), offers a first, potentially effective and efficient solution to address this problem.

Our findings suggest that the required main components for such a systems are sensors, pre-processing techniques, a machine learning algorithm, and actuators. The workflow of such systems follow the reactive agent model, where information about the environment is extracted and transformed into data by sensors, processed into input data for the machine learning algorithm, and then analyzed by the ML model. The outputs of the ML model which are realized by actuators and passed on as actions influence the environment.

The first step in setting up such a system should be considering in what environment this system will be deployed and the pre-requisites that the ML algorithm will need to account for when making its prediction. The environment heavily influences the choice of a suitable approach. There are mainly two approaches that can be followed: passive acoustic monitoring and passive visual monitoring. Usually visual sensors are preferred because they can concentrate various information in a single image and are thus very efficient. However, if there is the possibility that the image quality may be too low for the task at hand or that other environmental influences such as obstructions prevent images from being useful data [16], audio-based detection might be a suitable solution in individual use cases. Although, there are approaches that employ both techniques simultaneously, their results generally do not yield significantly better performances [73].

The sensor data type (i.e., audio or visual), however, does not influence the choice of machine learning algorithm. It is customary to transform every input data type into a visual representation via Mel spectrograms, followed by the application of an image detection ML algorithm on the processed data. The most preferred choice due to high performance is to use a pre-trained CNN model (such as VGG-16, VGG-18, and ResNet-50) via transfer learning, instead of building a custom model from scratch. However, the input data type influences on what type of data the ML model needs to be trained (i.e., does it need to recognize pictures of birds or frequency spectrograms). The data which is used for this training is usually pre-processed according to its signal-to-noise ration to obtain "clean" data which the algorithm can process. Some approaches consider training the algorithm also with noise data so that it can later differentiate between object and noise which improves results.

If the model has been trained sufficiently, its generated classifier rules can then be used to detect and classify birds in the real environment with a confident to high accuracy (approx. 80% − 90%), see Table 13.

This general workflow represents a simple but yet powerful application of machine learning when combined with actions that are derived from the output of the model. Once the algorithms have been trained, they can then be used to monitor the surrounding area for bird activity. If birds are detected in the vicinity of a wind turbine, the machine learning system can then trigger countermeasures, such as stopping the turbine within a short amount of time or redirecting the birds away from the blades through the use of deterrent mechanisms such as noise or light pulses [38], [39]. The key advantages of using machine learning for this use-case is that it can be performed in real-time, meaning that countermeasures can be activated before the birds reaches the dangerous area of the blades. This significantly increases the chances of avoiding collisions and reduces the risk to avifauna [16], [41].

The countermeasures could incorporate

- Turbine shutdown: The machine learning system can trigger the shutdown of a wind turbine when birds are detected in its vicinity. This prevents the birds from flying into the blades, getting hit by them and therefore eliminates the risk of collision.
- Acoustic deterrents: Acoustic deterrents are devices that emit high-pitched sounds that are designed to scare away birds. These deterrents can be triggered by the machine learning system when birds are detected in the vicinity of the wind turbines.
- Visual deterrents: Visual deterrents are devices that use flashing lights or bright colors to scare away birds. These deterrents can also be triggered by the machine learning system when birds are detected in the vicinity of the wind turbines [16].

In summary, machine learning models for bird detection paired with countermeasures provide a way to balance the tradeoff between expansion of wind farms and the preservation of avifauna by detecting birds and flying animals before they reach the danger zone of rotor blades and initiating counter measures to prevent collisions. This application of machine learning, therefore, may offer a way to contribute to the energy transition, prevent deadly interventions into the nature. In contrast to nuclear and coal power generation, the impact of wind energy is disproportionate, despite the high level of criticism. Regardless of whether total fatalities or mortalities relative to the power produced are compared, over 14 times as many avian individuals are killed through fossil fueled power generation, without exploitation of bird detection systems. The integration of such systems should further increase the difference and highlight the environmental compatibility of wind energy in contrast to previous fossil-fuel dominated generation portfolios [108].

B. CHALLENGES AND LIMITATIONS

As regards the completeness of analyzed papers it is striking that the use of sensors is a scarcely presented information. Often, the input data is based on given data sets recorded in the past, which lacks comparability regarding the performance of different sensors and neglects technical sensor improvements over time. Additionally, several works reduce their analysis on high-quality data which prevents an interpretation of the sensor's performance within realistic use cases. Hence, in combination with used input data it remains unclear, whether an environmental recording of audio is suitable to combine with high-volume sounds that are usually emitted by wind turbines and air planes. Therefore, a remaining question is as to whether technical noises would disturb the collection of data and thus the acoustic classification of birds in the specific use case, for instance within wind parks.

Another aspect that must be considered in the compatibility of acoustic detection in the wind energy sector use case is whether this approach can be implemented at wind turbines at all. The articles reviewed are largely related to habitat assessment, including ground-level fields or forests, for example, with good opportunities for sensor placement in a task-appropriate manner. However, wind turbines are usually located in a dispositional manner, birds that reach the height of the rotor blades and enter the flapping zone of wind turbines often have a different behaviour. The avifauna often does not stay stationary in this area, but rather crosses it when hunting or flying a longer distance [39]. It is unknown whether the birds emit enough vocalisations (songs, calls) during flight to be able detect a signal. In addition, there might exist strong variations between species if and when they sing or emit a warning call [90]. For example songbirds emit significantly more vocalisations than other species like birds of prey. Whose population is more sensible to fatalities due to smaller reproduction rates [38]. Thus, possibly only very inefficient acoustic approaches could be implemented in the wind sector, with high accuracy variability between classes. Moreover, birds must be detected in sufficient time, i.e. from a certain distance, to be able to activate the deterrent measurements safely [16]. In addition, acoustic systems have a weaker scalability compared to visual systems. Due to regional dialects, models may have to be adapted to the region of operation and trained in order to function efficiently, as reality may deviate from the trained laboratory condition of the algorithm [101], [102], [103]. In contrast, visual semantic components vary with season, gender and age, but usually remain the same within a class [76].

Therefore, although audio detection is on of the most prominent approaches of bird classification and may have proved its performance, this might be nevertheless an unsuitable approach for the targeted protection of many bird species in areal zones around wind rotor blade zones. In addition, bird species usually are located to their specific home habitats, which raises doubts whether acoustic learning allows for the global scaling for diverse species [101]. Thus, it is unclear, whether a visual approach might be better suitable for the detection of a bird in general (e.g. due to its wingspan), without the limitation of the detection of an unknown bird sound that prevents the assignment to its correct cause. Also, seasonal habits of specific birds, such as migrating species, are not specifically considered in the analyzed papers [39].

Another issue is the compilation of different performance indicators across all papers. Although there are several papers that have selected similar approaches to measure the accuracy of bird classification, many have chose distinct metrics [42]. The lack of comparability prevents an performance analysis across papers and thus the identification of the most promising bird detection systems within wind parks. Thus, an important question is, on what performance indicators researchers and practitioners want to agree on. Also other indicators as computational times are neglected so far. The needed time to classify an object is crucial to activate the deterrent measurements to prevent a collision accordingly. This question becomes even more relevant when considering political gatekeepers for the implementation of corresponding ML approaches. Only commonly defined performance indicators enable a goal-oriented work on specific thresholds and accumulate academic potentials for the improvement of accuracy rates and computational times [12], [16].

Although many approaches do not yet produce sufficiently robust classification results, the majority of the reviewed papers addresses the complex application of classifying avifauna explicitly into species. The sole detection of avifauna to activate deterrent mechanisms is less complex and already achieves accuracies of up to 100% [88]. A less advanced stage would be the categorisation of different birds according to their size. Though a consensus would have to be established on which birds are worthy of protection and to what extent, and which deterrent mechanisms are the most efficient for them. It would furthermore make the development and implementation of such systems easier if standards would be achieved of which mechanisms may be implemented under what conditions [12], [16], [38].

Similarly to the little content provided about sensors, the depth of actuators and actions as regards this review's aim is barely given. Within the analyzed papers there are few proposed systems to measure the evolution of bio diversity and perhaps even the contradictory effects of wind turbines, but solutions are still lacking to avoid corresponding damage on avifauna, thus combining the interests of ornithology and the energy transition. Although the risks and dangers are known and mentioned partially, there is no approach providing specific actions such as shutting down and restarting wind turbines, or the emission of acoustic deflection signals. Hence, a usable implementation in areas where birds are endangered is still not provided. Nevertheless, the measure of the evolution of bio diversity needs to be expanded since there is still no common data base regarding real impacts of wind turbines on avifauna [12].

Other issues are completely neglected such as the invasive influence of wind turbines in general. Wind turbines change environmental landscapes and therefore habitats. Some predators in particular utilize the area around wind turbines. The surfaces around hubs is usually flat and trimmed, therefore prey is nicely visible for the predators and might form a great hunting opportunity. In the past, it has been observed that birds of prey in particular can learn to recognise, circumvent and exploit wind turbines which might lead to a increased accurance of some birds around wind parks [39]. In contrast, birds of prey in particular may be caught by wind turbines [38], as their natural head shape means that they do not look forwards when flying, but downwards and therefore do not see the turbines [12], [16].

C. FUTURE RESEARCH OPPORTUNITIES

Based on the challenges and limitations there are several opportunities for future research:

- **Realistic field studies:** Studies so far limit their research mostly to pure machine learning processes. Realistic settings, including the test of different sensors and considering the disturbing noise influence of wind turbines are needed to allow interpretations regarding the suitability of a holistic system. Specific use cases, such as the attachment of sensors on wind turbines further allow to derive conclusions on different system ingredients, such as the used input data, pre-processing methods and ML approaches.
- **Commonly accepted performance indicators:** Diverse performance indicators prevent the possibility to compare different studies so far. Researchers and practitioners need to agree on the most suitable performance indicators in combination with wind parks accordingly. In a transit phase, this possibly raises the need to contrast different approaches and debate about benefits and disadvantages of these indicators. Other indicators then need to be implemented, such as computational times, that not only measure the pure accuracy, but further influence real collision outcomes. Strong performance indicators as well as the presentation of a working, holistic system are crucial to persuade political gatekeepers.
- **Academic collaboration with biology sciences:** Many biologic questions are still unclear, such as when and why specific bird species communicate, the consequences of migrating birds and the behavior of birds in the vicinity of wind parks. A deep knowledge of biologic and anatomic states is needed to understand bird behavior and derive countermeasures of endangered birds in the vicinity of wind turbines. Hence, biologic know-know should be integrated to discuss avoidance strategies that are both, technically effective and biologically appropriate.
- **Practical cooperations:** Although there are many approaches that are able to assess the evolution of avifauna as an outcome of possible collisions, avoidance

strategies are not tested so far. In combination with a common approach to measure bird populations, different opportunities such as the shut-down of wind turbines, acoustic and visual deterrents need to be tested in corporation with wind turbine managers to measure real-setting effects of these avoidance strategies. Possible unknown (and maybe invasive) side effects of those strategies as well as economic and biological outcomes then can be estimated.

VI. CONCLUSION

In light of the increasing global impacts of climate change, it is necessary to keep up with the energy transition and its goals to move towards more emission-neutral and renewable energies. Thereby, it is imperative to make use of all types of green energy sources that are available in order to keep up with the growing demand for energy from industries and private households.

Wind energy is a rapidly growing source of energy, and next to hydropower and solar energy, the most commonly used form of renewable and sustainable energy generation. Nonetheless, the expansion of wind farms in regions with significant wind energy potential has stalled due to rightful concerns about adverse impacts on the avifauna. For instance, birds are impacted by the rapidly rotating blades of wind turbines. This circumstance leads to a paradoxical trade-off between facilitating the energy transition, whose goal is to preserve our environment, and the detrimental effects on the same environment it seeks to preserve.

This trade-off between the expansion of renewable energies and adverse effects on the environment necessitates a way to detect flying animals before they reach the danger zone of the rotor blades and to develop appropriate measures to deter them from flying into this area and, subsequently, prevent collisions. Such approaches should work automatically and with high reliability, which implies the use of automated machines that can recognize birds and initiate countermeasures in parallel. In this regard, machine learning applications might constitute a viable way to resolve this problem. However, no work has systematically analyzed this solution approach yet. Thus, knowledge and insights about design choices, advantages, and limitations are scarce.

Consequently, our research objective was to analyze the body of informatics-related literature in this area of research. We found various works that cover machine learning algorithms for detecting and classifying birds and bats. Although some papers consider the possibility of using these technical artifacts in an application, such as detecting and preventing bird collusion in wind parks, none implement and test such an application. Nevertheless, machine learning algorithms offer the desired properties of high automation in combination with high accuracy, leaving plenty of room for further work to design systems around these artifacts.

While the algorithms are essential for the solution approach, more components need to be considered for a working system that enables this approach. The reactive agent model provides a common framework for components that machines need to interact with the environment in which they have been deployed and in which they should perform their task. Besides the machine learning algorithms, such a system needs sensors and actuators to receive perceptions from the environment and execute follow-up actions. These two system components are currently significantly understudied and require more attention from scholars and practitioners alike.

We, therefore, conclude this work with a call for action. In order to facilitate the energy transition effectively, we need to make the most of the renewable energy sources available to our society while still preserving the environment, which is the foundation of our very existence. Machine learning-based approaches can help to achieve this goal; however, further work and insights from the surrounding components (i.e., sensors and actuators) are needed in this context. The research gaps that have been revealed demand collaboration between specialists in sensors, actuators, and machine learning algorithms, as well as wind turbine operators, policymakers, ornithologists, and other stakeholders in a trans-disciplinary manner to shed light on the journey from our current fossil fuel-dependent society to a future that is more sustainable.

REFERENCES

[1] J. Barnett and W. N. Adger, "Climate change, human security and violent conflict," *Political Geography*, vol. 26, no. 6, pp. 639–655, Aug. 2007.

[2] B. C. O'Neill, M. Oppenheimer, R. Warren, S. Hallegatte, R. E. Kopp, H. O. Pörtner, R. Scholes, J. Birkmann, W. Foden, R. Licker, K. J. Mach, P. Marbaix, M. D. Mastrandrea, J. Price, K. Takahashi, J.-P. Van Ypersele, and G. Yohe, "IPCC reasons for concern regarding climate change risks," *Nature Climate Change*, vol. 7, no. 1, pp. 28–37, Jan. 2017.

[3] L. Bernstein, P. Bosch, O. Canziani, Z. Chen, R. Christ, and K. Riahi, "IPCC, 2007: Climate change 2007: Synthesis report," Contribution Work. Groups I, II III Fourth Assessment Rep. Intergovernmental Panel Climate Change, Core Writing Team, R. K. Pachauri and A. Reisinger, Eds., IPCC, Geneva, Switzerland, p. 104. [Online]. Available: https://www.ipcc.ch/report/ar4/syr/

[4] R. S. J. Tol, "The economic effects of climate change," *J. Econ. Perspect.*, vol. 23, no. 2, pp. 29–51, Apr. 2009.

[5] United Nations, "The Paris agreement—Publication," in *Proc. Paris Climate Change Conf.*, Nov. 2015, pp. 1–54. [Online]. Available: https://unfccc.int/documents/184656

[6] M. Karmellos, D. Kopidou, and D. Diakoulaki, "A decomposition analysis of the driving factors of CO_2 (carbon dioxide) emissions from the power sector in the European union countries," *Energy*, vol. 94, pp. 680–692, Jan. 2016.

[7] M. Daroń and M. Wilk, "Management of energy sources and the development potential in the energy production sector—A comparison of EU countries," *Energies*, vol. 14, no. 3, p. 685, Jan. 2021.

[8] T. Goh, B. W. Ang, B. Su, and H. Wang, "Drivers of stagnating global carbon intensity of electricity and the way forward," *Energy Policy*, vol. 113, pp. 149–156, Feb. 2018.

[9] R. Sims, "Renewable energy: A response to climate change," *Sol. Energy*, vol. 76, nos. 1–3, pp. 9–17, Jan. 2004.

[10] B. D. Solomon and K. Krishna, "The coming sustainable energy transition: History, strategies, and outlook," *Energy Policy*, vol. 39, no. 11, pp. 7422–7431, Nov. 2011.

[11] S. Wang, S. Wang, and P. Smith, "Ecological impacts of wind farms on birds: Questions, hypotheses, and research needs," *Renew. Sustain. Energy Rev.*, vol. 44, pp. 599–607, Apr. 2015.

[12] A. T. Marques, H. Batalha, S. Rodrigues, H. Costa, M. J. R. Pereira, C. Fonseca, M. Mascarenhas, and J. Bernardino, "Understanding bird collisions at wind farms: An updated review on the causes and possible mitigation strategies," *Biol. Conservation*, vol. 179, pp. 40–52, Nov. 2014.

[13] P. Plonczkier and I. C. Simms, "Radar monitoring of migrating pink-footed geese: Behavioural responses to offshore wind farm development," *J. Appl. Ecology*, vol. 49, no. 5, pp. 1187–1194, Oct. 2012, doi: 10.1111/j.1365-2664.2012.02181.x.

[14] W. P. Erickson, G. D. Johnson, D. M. Strickland, D. P. Young Jr., K. J. Sernka, and R. E. Good, "Avian collisions with wind turbines: A summary of existing studies and comparisons to other sources of avian collision mortality in the United States," Western EcoSystems Technol., Inc., Cheyenne, WY, USA, Tech. Rep., DOE-00SF22100, 2001.

[15] K. S. Smallwood and D. A. Bell, "Effects of wind turbine curtailment on bird and bat fatalities," *J. Wildlife Manage.*, vol. 84, no. 4, pp. 685–696, May 2020.

[16] D. Gradolewski, D. Dziak, M. Martynow, D. Kaniecki, A. Szurlej-Kielanska, A. Jaworski, and W. J. Kulesza, "Comprehensive bird preservation at wind farms," *Sensors*, vol. 21, no. 1, p. 267, Jan. 2021.

[17] G. Maclaurin, C. Hein, T. Williams, O. Roberts, E. Lantz, G. Buster, and A. Lopez, "National–scale impacts on wind energy production under curtailment scenarios to reduce bat fatalities," *Wind Energy*, vol. 25, no. 9, pp. 1514–1529, 2022.

[18] M. de Lucas, M. Ferrer, M. J. Bechard, and A. R. Muñoz, "Griffon vulture mortality at wind farms in southern Spain: Distribution of fatalities and active mitigation measures," *Biol. Conservation*, vol. 147, no. 1, pp. 184–189, Mar. 2012.

[19] E. B. Arnett and R. F. May, "Mitigating wind energy impacts on wildlife: Approaches for multiple taxa," *Hum.-Wildlife Interact.*, vol. 10, no. 1, p. 5, 2016.

[20] P. P. Shinde and S. Shah, "A review of machine learning and deep learning applications," in *Proc. 4th Int. Conf. Comput. Commun. Control Autom. (ICCUBEA)*, Pune, India, Aug. 2018, pp. 1–6.

[21] B. K. Bose, "Artificial intelligence techniques in smart grid and renewable energy systems—Some example applications," *Proc. IEEE*, vol. 105, no. 11, pp. 2262–2273, Nov. 2017.

[22] T. Bhardwaj, S. Mehenge, and B. S. Revathi, "Wind turbine power output forecasting using artificial intelligence," in *Proc. Int. Virtual Conf. Power Eng. Comput. Control: Develop. Electr. Vehicles Energy Sector Sustain. Future (PECCON)*, May 2022, pp. 1–5.

[23] A. Višković, V. Franki, and D. Jevtic, "Artificial intelligence as a facilitator of the energy transition," in *Proc. 45th Jubilee Int. Conv. Inf., Commun. Electron. Technol. (MIPRO)*, May 2022, pp. 494–499.

[24] J. T. Dellosa and E. C. Palconit, "Artificial intelligence (AI) in renewable energy systems: A condensed review of its applications and techniques," in *Proc. IEEE Int. Conf. Environ. Electr. Eng. IEEE Ind. Commercial Power Syst. Eur. (EEEIC I&CPS Europe)*, Sep. 2021, pp. 1–6.

[25] P. Datar, K. Jain, and B. Dhedhi, "Detection of birds in the wild using deep learning methods," in *Proc. 4th Int. Conf. Converg. Technol. (I2CT)*, Oct. 2018, pp. 1–4.

[26] X. Wei, Y. Song, O. M. Aodha, J. Wu, Y. Peng, J. Tang, J. Yang, and S. Belongie, "Fine-grained image analysis with deep learning: A survey," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 44, no. 12, pp. 8927–8948, Dec. 2022.

[27] J. Wäldchen and P. Mäder, "Machine learning for image based species identification," *Methods Ecology Evol.*, vol. 9, no. 11, pp. 2216–2225, Nov. 2018.

[28] W. Thuiller, "Climate change and the ecologist," *Nature*, vol. 448, no. 7153, pp. 550–552, Aug. 2007.

[29] A. Arneth, Y.-J. Shin, P. Leadley, C. Rondinini, E. Bukvareva, M. Kolb, G. F. Midgley, T. Oberdorff, I. Palomo, and O. Saito, "Post-2020 biodiversity targets need to embrace climate change," *Proc. Nat. Acad. Sci. USA*, vol. 117, no. 49, pp. 30882–30891, Dec. 2020.

[30] S. T. Fakana, "Causes of climate change," *Glob. J. Sci. Front Res., H Environ. Earth Sci.*, vol. 20, no. 2, pp. 7–12, 2020.

[31] A. S. Mori, L. E. Dee, A. Gonzalez, H. Ohashi, J. Cowles, A. J. Wright, M. Loreau, Y. Hautier, T. Newbold, P. B. Reich, T. Matsui, W. Takeuchi, K.-I. Okada, R. Seidl, and F. Isbell, "Biodiversity–productivity relationships are key to nature-based climate solutions," *Nature Climate Change*, vol. 11, no. 6, pp. 543–550, Jun. 2021.

[32] R. York and S. E. Bell, "Energy transitions or additions?: Why a transition from fossil fuels requires more than the growth of renewable energy," *Energy Res. Social Sci.*, vol. 51, pp. 40–43, 2019.

[33] A. Qazi, F. Hussain, N. ABD. Rahim, G. Hardaker, D. Alghazzawi, K. Shaban, and K. Haruna, "Towards sustainable energy: A systematic review of renewable energy sources, technologies, and public opinions," *IEEE Access*, vol. 7, pp. 63837–63851, 2019.

[34] IE Agency. (2020). *Europe Data Explorer: Renewable Electricity Generation by Source (Non-Combustible), Europe 1990–2020*. [Online]. Available: https://www.iea.org/regions/europe

[35] N. Hatziargyriou and A. Zervos, "Wind power development in Europe," *Proc. IEEE*, vol. 89, no. 12, pp. 1765–1782, Oct. 2001.

[36] European Commission. (2021). *Fit for 55: Delivering the Eu's 2030 Climate Target on the Way to Climate Neutrality*. [Online]. Available: https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A52021DC0550

[37] European Commission. (2020). *EU Biodiversity Strategy for 2030*. [Online]. Available: https://eur-lex.europa.eu/legal-content/EN/TXT/?qid=1590574123338&uri=CELEX%3A52020DC0380

[38] C. Heuck, C. Herrmann, C. Levers, P. J. Leitão, O. Krone, R. Brandl, and J. Albrecht, "Wind turbines in high quality habitat cause disproportionate increases in collision mortality of the white-tailed eagle," *Biol. Conservation*, vol. 236, pp. 44–51, Aug. 2019.

[39] J. L. Tellería, "Potential impacts of wind farms on migratory birds crossing Spain," *Bird Conservation Int.*, vol. 19, no. 2, pp. 131–136, Jun. 2009.

[40] Avibase. (2022). *Avibase—The World Bird Database*. [Online]. Available: https://avibase.bsc-eoc.org/avibase.jsp?lang=EN

[41] R. M. R. Barclay, E. F. Baerwald, and J. C. Gruver, "Variation in bat and bird fatalities at wind energy facilities: Assessing the effects of rotor size and tower height," *Can. J. Zool.*, vol. 85, no. 3, pp. 381–387, Feb. 2007.

[42] J. Xie, Y. Zhong, J. Zhang, S. Liu, C. Ding, and A. Triantafyllopoulos, "A review of automatic recognition technology for bird vocalizations in the deep learning era," *Ecol. Informat.*, vol. 73, Mar. 2023, Art. no. 101927.

[43] K. Nagy, T. Cinkler, C. Simon, and R. Vida, "Internet of birds (IoB): Song based bird sensing via machine learning in the cloud : How to sense, identify, classify birds based on their songs?" in *Proc. IEEE SENSORS*, Oct. 2020, pp. 1–4.

[44] S. J. Russell and P. Norvig, *Artificial Intelligence a Modern Approach*. London, U.K.: Pearson Education, 2010.

[45] L. Zhou, S. Pan, J. Wang, and A. V. Vasilakos, "Machine learning on big data: Opportunities and challenges," *Neurocomputing*, vol. 237, pp. 350–361, May 2017.

[46] S. Nolfi, "Power and the limits of reactive agents," *Neurocomputing*, vol. 42, nos. 1–4, pp. 119–145, Jan. 2002.

[47] B. Dafflon, J. Contet, F. Gechter, and P. Gruer, "Toward a reactive agent based parking assistance system," in *Proc. IEEE 24th Int. Conf. Tools Artif. Intell.*, vol. 1, Nov. 2012, pp. 500–507.

[48] B. Dafflon, F. Gechter, P. Gruer, and A. Koukam, "Vehicle platoon and obstacle avoidance: A reactive agent approach," *IET Intell. Transp. Syst.*, vol. 7, no. 3, pp. 257–264, Sep. 2013.

[49] A. Milani and V. Poggioni, "Planning in reactive environments," *Comput. Intell.*, vol. 23, no. 4, pp. 439–463, Dec. 2007.

[50] L.-J. Lin, "Self-improving reactive agents based on reinforcement learning, planning and teaching," *Mach. Learn.*, vol. 8, nos. 3–4, pp. 293–321, May 1992.

[51] T. P. Carvalho, F. A. A. M. N. Soares, R. Vita, R. P. Da Francisco, J. P. Basto, and S. G. S. Alcalá, "A systematic literature review of machine learning methods applied to predictive maintenance," *Comput. Ind. Eng.*, vol. 137, pp. 1–10, Nov. 2019. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0360835219304838

[52] S. Duda, D. Geyer, T. Guggenberger, M. Principato, and D. Protschky, "A systematic literature review on how to improve the privacy of artificial intelligence using blockchain," in *Proc. Pacific Asia Conf. Inf. Syst.*, 2022, pp. 1–13. [Online]. Available: https://aisel.aisnet.org/pacis2022/176/

[53] R. Sharma, S. S. Kamble, A. Gunasekaran, V. Kumar, and A. Kumar, "A systematic literature review on machine learning applications for sustainable agriculture supply chain performance," *Comput. Oper. Res.*, vol. 119, pp. 1–17, Jul. 2020.

[54] Y. Levy and T. J. Ellis, "A systems approach to conduct an effective literature review in support of information systems research," *Informing Sci., Int. J. Emerg. Transdiscipline*, vol. 9, pp. 181–212, Jan. 2006.

[55] J. Vom Brocke, A. Simons, K. Riemer, B. Niehaves, R. Plattfaut, and A. Cleven, "Standing on the shoulders of giants: Challenges and recommendations of literature search in information systems research," *Commun. Assoc. Inf. Syst.*, vol. 37, no. 1, p. 9, 2015.

[56] H. Snyder, "Literature review as a research methodology: An overview and guidelines," *J. Bus. Res.*, vol. 104, pp. 333–339, Nov. 2019.

[57] B. Kitchenham and S. Charters, "Guidelines for performing systematic literature reviews in software engineering," Keele Univ., Durham Univ. Joint Rep., Tech. Rep. EBSE 2007-001, 2007.

[58] J. Webster and R. T. Watson, "Analyzing the past to prepare for the future: Writing a literature review," *MIS Quart.*, vol. 26, no. 2, pp. 13–23, 2002.

[59] H. Zhang, M. A. Babar, and P. Tell, "Identifying relevant studies in software engineering," *Inf. Softw. Technol.*, vol. 53, no. 6, pp. 625–637, Jun. 2011.

[60] T. Dyba, T. Dingsoyr, and G. K. Hanssen, "Applying systematic reviews to diverse study types: An experience report," in *Proc. 1st Int. Symp. Empirical Softw. Eng. Meas. (ESEM)*, Sep. 2007, pp. 225–234.

[61] L. Chen, M. A. Babar, and H. Zhang, "Towards an evidence-based understanding of electronic data sources," in *Proc. 14th Int. Conf. Eval. Assessment Softw. Eng.*, Swindon, U.K.: BCS Learning & Development Ltd., 2010, pp. 135–138.

[62] J. Cohen, "A coefficient of agreement for nominal scales," *Educ. Psychol. Meas.*, vol. 20, no. 1, pp. 37–46, Apr. 1960.

[63] L. J. Cronbach, "Coefficient alpha and the internal structure of tests," *Psychometrika*, vol. 16, no. 3, pp. 297–334, Sep. 1951.

[64] K. Krippendorff, "Reliability in content analysis: Some common misconceptions and recommendations," *Hum. Commun. Res.*, vol. 30, no. 3, pp. 411–433, Jul. 2004.

[65] M. Lombard, J. Snyder-Duch, and C. C. Bracken, "Content analysis in mass communication: Assessment and reporting of intercoder reliability," *Human Commun. Res.*, vol. 28, no. 4, pp. 587–604, Oct. 2002.

[66] K. Krippendorff, "Computing Krippendorff's alpha-reliability," Tech. Rep. 1750, 2011, pp. 1–10. [Online]. Available: https://repository.upenn.edu/asc_papers/43

[67] J.-Y. Antoine, J. Villaneau, and A. Lefeuvre, "Weighted Krippendorff's alpha is a more reliable metrics for multi-coders ordinal annotations: Experimental studies on emotion, opinion and coreference annotation," in *Proc. 14th Conf. Eur. Chapter Assoc. Comput. Linguistics*, Gothenburg, Sweden, 2014, pp. 550–559. [Online]. Available: https://aclanthology.org/E14-1058

[68] K. L. Gwet, "On the Krippendorff's alpha coefficient,"Adv. Anal., Gaithersburg, MD, USA, Tech. Rep. 1758, 2011, pp. 1–16. [Online]. Available: https://agreestat.com/papers/onkrippendorffalpha_rev10052015.pdf

[69] C. Wohlin, P. Runeson, M. Höst, M. C. Ohlsson, B. Regnell, and A. Wesslén, *Experimentation in Software Engineering*. Berlin, Germany: Heidelberg: Springer, 2012.

[70] N. Yoh, T. Kingston, E. McArthur, O. E. Aylen, J. C. C. Huang, E. R. Jinggong, F. A. A. Khan, B. P. Lee, S. L. Mitchell, J. E. Bicknell, and M. J. Struebig, "A machine learning framework to classify Southeast Asian echolocating bats," *Ecol. Indicators*, vol. 136, Jan. 2022, Art. no. 108696.

[71] J. LeBien, M. Zhong, M. Campos-Cerqueira, J. P. Velev, R. Dodhia, J. L. Ferres, and T. M. Aide, "A pipeline for identification of bird and frog species in tropical soundscape recordings using a convolutional neural network," *Ecol. Informat.*, vol. 59, Sep. 2020, Art. no. 101113.

[72] B. A. K. S. G. V. Chandu Munikoti Murthy Murthy and C. Nagaraj, "Automated bird species identification using audio signal processing and neural networks, Amaravati, India," in *Proc. Int. Conf. Artif. Intell. Signal Process. (AISP)*, 2020, pp. 1–5.

[73] N. Sharma, A. Vijayeendra, V. Gopakumar, P. Patni, and A. Bhat, "Automatic identification of bird species using Audio/Video processing," in *Proc. Int. Conf. Advancement Technol. (ICONAT)*, Goa, India, Jan. 2022, pp. 1–6.

[74] T. L. F. Evangelista, T. M. Priolli, C. N. Silla, B. A. Angelico, and C. A. A. Kaestner, "Automatic segmentation of audio signals for bird species identification," in *Proc. IEEE Int. Symp. Multimedia*, Taichung, Taiwan, Dec. 2014, pp. 223–228.

[75] Á. Incze, H. Jancsó, Z. Szilágyi, A. Farkas, and C. Sulyok, "Bird sound recognition using a convolutional neural network," in *Proc. IEEE 16th Int. Symp. Intell. Syst. Informat. (SISY)*, Subotica, Serbia, Sep. 2018, pp. 000295–000300.

[76] A. Marini, J. Facon, and A. L. Koerich, "Bird species classification based on color features," in *Proc. IEEE Int. Conf. Syst., Man, Cybern.*, Manchester, U.K., Oct. 2013, pp. 4336–4341.

[77] S. Islam, S. I. A. Khan, M. M. Abedin, K. M. Habibullah, and A. K. Das, "Bird species classification from an image using VGG-16 network," in *Proc. 7th Int. Conf. Comput. Commun. Manage.*, Bangkok, Thailand, Jul. 2019, pp. 38–42.

[78] N. Bold, C. Zhang, and T. Akashi, "Bird species classification with audio-visual data using CNN and multiple kernel learning," in *Proc. Int. Conf. Cyberworlds (CW)*, Oct. 2019, pp. 85–88.

[79] S. Kahl, C. M. Wood, M. Eibl, and H. Klinck, "BirdNET: A deep learning solution for avian diversity monitoring," *Ecol. Informat.*, vol. 61, pp. 101236–101245, Mar. 2021.

[80] S. Aggarwal and S. Sehgal, "Classification of bird species using audio processing and deep neural network," in *Proc. 3rd Int. Conf. Intell. Comput. Instrum. Control Technol. (ICICICT)*, Kannur, India, Aug. 2022, pp. 138–143.

[81] A. A. Hidayat, T. W. Cenggoro, and B. Pardamean, "Convolutional neural networks for scops owl sound classification," *Proc. Comput. Sci.*, vol. 179, pp. 81–87, Jan. 2021.

[82] F. Yang, Y. Jiang, and Y. Xu, "Design of bird sound recognition model based on lightweight," *IEEE Access*, vol. 10, pp. 85189–85198, 2022.

[83] M. T. Lopes, A. L. Koerich, C. N. Silla, and C. A. A. Kaestner, "Feature set comparison for automatic bird species identification," in *Proc. IEEE Int. Conf. Syst., Man, Cybern.*, Anchorage, AK, USA, Oct. 2011, pp. 965–970.

[84] S. Bhusal, U. Bhattarai, and M. Karkee, "Improving pest bird detection in a vineyard environment using super-resolution and deep learning," *IFAC-PapersOnLine*, vol. 52, no. 30, pp. 18–23, 2019.

[85] Y. Jadhav, V. Patil, and D. Parasar, "Machine learning approach to classify birds on the basis of their sound," in *Proc. Int. Conf. Inventive Comput. Technol. (ICICT)*, Feb. 2020, pp. 69–73.

[86] L. B. Boudaoud, F. Maussang, R. Garello, and A. Chevallier, "Marine bird detection based on deep learning using high-resolution aerial images," in *Proc. OCEANS*, Marseille, France, Jun. 2019, pp. 1–7.

[87] K. M. Ragib, R. T. Shithi, S. A. Haq, M. Hasan, K. M. Sakib, and T. Farah, "PakhiChini: Automatic bird species identification using deep learning," in *Proc. 4th World Conf. Smart Trends Syst., Secur. Sustainability (WorldS4)*, London, U.K., Jul. 2020, pp. 1–6.

[88] Y. Huang and H. Basanta, "Recognition of endemic bird species using deep learning models," *IEEE Access*, vol. 9, pp. 102975–102984, 2021.

[89] R. Kojima, O. Sugiyama, R. Suzuki, K. Nakadai, and C. E. Taylor, "Semi-automatic bird song analysis by spatial-cue-based integration of sound source detection, localization, separation, and identification," in *Proc. IEEE/RSJ Int. Conf. Intell. Robots Syst. (IROS)*, Daejeon, South Korea, Oct. 2016, pp. 1287–1292.

[90] R. P. Tivarekar and H. G. Virani, "Species recognition using audio processing algorithm," in *Proc. Int. Conf. Trends Electron. Informat. (ICEI)*, Tirunelveli, India, May 2017, pp. 527–532.

[91] Y. Li, H. Zhou, and Y. Zhao, "The fine-grained recognition of bird images based on joint semantic components and ResNet," in *Proc. 5th Int. Conf. Artif. Intell. Big Data (ICAIBD)*, Chengdu, China, May 2022, pp. 406–410.

[92] Z. J. Ruff, D. B. Lesmeister, C. L. Appel, and C. M. Sullivan, "Workflow and convolutional neural network for automated identification of animal sounds," *Ecol. Indicators*, vol. 124, pp. 107419–107430, Oct. 2021.

[93] D. Stowell and M. D. Plumbley, "Automatic large-scale classification of bird sounds is strongly improved by unsupervised feature learning," *PeerJ*, vol. 2, pp. 1–31, Jul. 2014.

[94] C.-L. Yang, Y. Harjoseputro, Y.-C. Hu, and Y.-Y. Chen, "An improved transfer-learning for image-based species classification of protected Indonesians birds," *Comput., Mater. Continua*, vol. 73, no. 3, pp. 4577–4593, 2022.

[95] D. Kaminska and A. Gmerek, "Automatic identification of bird species: A comparison between kNN and SOM classifiers," in *Proc. Joint Conf. New Trends Audio Video Signal Process., Algorithms, Archit., Arrangements Appl. (NTAV/SPA)*, Lodz, Poland, Sep. 2012, pp. 77–82.

[96] R. A. Bistel, A. Martinez, and G. B. Mindlin, "Neural networks that locate and identify birds through their songs," *Eur. Phys. J. Special Topics*, vol. 231, no. 3, pp. 185–194, Apr. 2022.

[97] G. Gupta, M. Kshirsagar, M. Zhong, S. Gholami, and J. L. Ferres, "Comparing recurrent convolutional neural networks for large scale bird species classification," *Sci. Rep.*, vol. 11, no. 1, pp. 1–12, Aug. 2021.

[98] B. Qiao, Z. Zhou, H. Yang, and J. Cao, "Bird species recognition based on SVM classifier and decision tree," in *Proc. 1st Int. Conf. Electron. Instrum. Inf. Syst. (EIIS)*, Harbin, China, Jun. 2017, pp. 1–4.

[99] T. Choudhary, S. Gujar, K. Panchal, V. Mishra, and A. Goswami, "A deep learning-based transfer learning approach for the bird species classification," in *Proc. 10th Int. Conf. Adv. Comput. (IACC)*. Goa, India: Springer, Dec. 2020, pp. 43–52.

[100] J. Verbraeken, M. Wolting, J. Katzy, J. Kloppenburg, T. Verbelen, and J. S. Rellermeyer, "A survey on distributed machine learning," *ACM Comput. Surv.*, vol. 53, no. 2, pp. 1–33, Mar. 2020.

[101] S. D. Hill, W. Ji, K. A. Parker, C. Amiot, and S. J. Wells, "A comparison of vocalisations between Mainland tui (Prosthemadera novaeseelandiae novaeseelandiae) and Chatham Island tui (P. n. chathamensis)," *New Zealand J. Ecol.*, vol. 37, no. 2, pp. 214–223, 2013.

[102] J. Podos and P. S. Warren, "The evolution of geographic variation in birdsong," *Adv. Study Behav.*, vol. 37, pp. 403–458, Jan. 2007.

[103] T. F. Wright, "Regional dialects in the contact call of a parrot," in *Proc. Roy. Soc. London. B, Biol. Sci.*, vol. 263, no. 1372, pp. 867–872, 1996.

[104] V. Powys, "Regional variation in the territorial songs of superb lyrebirds in the central tablelands of new South Wales," *Emu Austral Ornithology*, vol. 95, no. 4, pp. 280–289, Dec. 1995.

[105] V. Van Zalinge. (2015). *Selective Focus Photography of Blue Kingfisher*. Accessed: Feb. 19, 2023. [Online]. Available: https://unsplash.com/de/fotos/vUNQaTtZeOo

[106] S. A. Sanchez, H. J. Romero, and A. D. Morales, "A review: Comparison of performance metrics of pretrained models for object detection using the tensorflow framework," in *Proc. IOP Conf. Mater. Sci. Eng.*, vols. 844–860, no. 1, 2020, pp. 1–25.

[107] S. Davis and P. Mermelstein, "Comparison of parametric representations for monosyllabic word recognition in continuously spoken sentences," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. ASSP-28, no. 4, pp. 357–366, Aug. 1980.

[108] B. K. Sovacool, "Contextualizing avian mortality: A preliminary appraisal of bird and bat fatalities from wind, fossil-fuel, and nuclear electricity," *Energy Policy*, vol. 37, no. 6, pp. 2241–2248, Jun. 2009.

MARC PRINCIPATO was born in Heidelberg, Germany, in 2000. He received the B.Sc. degree in business administration with a major in information systems and information management from the University of Bayreuth, Bayreuth, Germany, in 2022. He is currently pursuing the M.Sc. degree in business administration with a specialization in technology, operations, and processes.

He is also a Research Assistant with the University of Bayreuth and the Branch Business & Information Systems Engineering, Fraunhofer Institute for Applied Information Technology (FIT), Bayreuth. He has coauthored "A Systematic Literature Review on How to Improve the Privacy of Artificial Intelligence Using Blockchain" (2022) (PACIS 2022 Proceedings), "Decentralized Finance—The Rise of a New Paradigm?" (2022) (Rethinking Finance) (vol. 6), and "A Multivocal Literature Review of Decentralized Finance: Current Knowledge and Future Research Avenues" (2023) of *Electronic Markets* (vol. 33). His research interests include blockchains and applied cryptography for data sovereignty and privacy in information systems.

LISA HASSELWANDER was born in Cologne, Germany, in 1996. She received the B.Sc. degree in business administration with a major in informatics systems and technology and information management from the University of Bayreuth, Germany, in 2022, where she is currently pursuing the M.Sc. degree in business administration with specializations in informatics systems and technology and information management.

During her studies, she was with the University of Graz, Austria, in 2022, she focused on sustainability management and energy and resource economics. From 2018 to 2019, she was a Research Assistant with the University of Bayreuth and supported the institutions office for sustainability in promoting sustainable development. In 2020, she completed an internship in renewable energy project management with Greek Consultancy Infratec SE, Athens, Greece, focused on solar and wind projects. Since 2021, she has been a working student with Kommunalkredit Public Consulting GmbH, Vienna, Austria, supporting the company in the processing of climate and environmental protection funding cases in the sectors home and energy.

MICHAEL STANGNER was born in Dorsten, Germany, in 1990. He received the B.Sc. degree in business administration with a major in marketing, and strategic management and organization from the University of Bayreuth, Bayreuth, Germany, in 2017, where he is currently pursuing the M.Sc. degree in business administration, with a specialization in marketing, and entrepreneurship, and digital business models. From 2010 to 2012, he was an Editorial Soldier in media battalion operative information with German Armed Forces, Koblenz, Germany. He supported the Continental AG, Hanover, Germany, in corporate communications, in 2016, and Silega as the Key Account Manager in Mexico City, Mexico, from 2017 to 2018. From 2019 to 2020, he was with Wilkinson Sword GmbH (Edgewell Personal Care), Solingen, Germany, most recently as an European Portfolio Coordinator. Since 2021, he has been with Kantar, Munich, Germany, in insights division and the domain of brand, media, and creative.

RICARDO BUETTNER (Senior Member, IEEE) received the Dipl.-Inf. degree in computer science and the Dipl.-Wirtsch.-Ing. degree in industrial engineering and management from the Technical University of Ilmenau, Germany, the Dipl.-Kfm. degree in business administration from the University of Hagen, Germany, the Ph.D. degree in information systems from the University of Hohenheim, Germany, and the Habilitation (venia legendi) degree in information systems from the University of Trier, Germany. He is currently a Chaired Professor in information systems and data science with the University of Bayreuth, Germany. He has published over 140 peer-reviewed articles, including articles in *Electronic Markets*, *AIS Transactions on Human–Computer Interaction*, *Personality and Individual Differences*, *European Journal of Psychological Assessment*, *PLOS One*, and IEEE Access. He has received 17 international best paper, a best reviewer, and the service awards and award nominations, including best paper awards by *AIS Transactions on Human–Computer Interaction*, *Electronic Markets*, and HICSS, for his work.