


Review

# Oil Palm and Machine Learning: Reviewing One Decade of Ideas, Innovations, Applications, and Gaps

Nuzhat Khan <sup>1</sup>, Mohamad Anuar Kamaruddin <sup>1,\*</sup>, Usman Ullah Sheikh <sup>2</sup>, Yusri Yusup <sup>1</sup>  
and Muhammad Paend Bakht <sup>2</sup>

<sup>1</sup> School of Industrial Technology, Universiti Sains Malaysia, Gelugor 11800, Malaysia; nuzhat\_khan@student.usm.my (N.K.); yusriy@usm.my (Y.Y.)

<sup>2</sup> School of Electrical Engineering, Universiti Teknologi Malaysia, Johor Bahru 81310, Malaysia; usman@fke.utm.my (U.U.S.); paend.bakht@graduate.utm.my (M.P.B.)

\* Correspondence: anuarkamaruddin@usm.my

**Abstract:** Machine learning (ML) offers new technologies in the precision agriculture domain with its intelligent algorithms and strong computation. Oil palm is one of the rich crops that is also emerging with modern technologies to meet global sustainability standards. This article presents a comprehensive review of research dedicated to the application of ML in the oil palm agricultural industry over the last decade (2011–2020). A systematic review was structured to answer seven predefined research questions by analysing 61 papers after applying exclusion criteria. The works analysed were categorized into two main groups: (1) regression analysis used to predict fruit yield, harvest time, oil yield, and seasonal impacts and (2) classification techniques to classify trees, fruit, disease levels, canopy, and land. Based on defined research questions, investigation of the reviewed literature included yearly distribution and geographical distribution of articles, highly adopted algorithms, input data, used features, and model performance evaluation criteria. Detailed quantitative–qualitative investigations have revealed that ML is still underutilised for predictive analysis of oil palm. However, smart systems integrated with machine vision and artificial intelligence are evolving to reform oil palm agri-business. This article offers an opportunity to understand the significance of ML in the oil palm agricultural industry and provides a roadmap for future research in this domain.

**Keywords:** oil palm; machine learning; systematic review; agriculture; sustainability



check for updates

**Citation:** Khan, N.; Kamaruddin, M.A.; Sheikh, U.U.; Yusup, Y.; Bakht, M.P. Oil Palm and Machine Learning: Reviewing One Decade of Ideas, Innovations, Applications, and Gaps. *Agriculture* **2021**, *11*, 832. <https://doi.org/10.3390/agriculture11090832>

Academic Editor: Yanbo Huang

Received: 12 July 2021

Accepted: 25 August 2021

Published: 31 August 2021

**Publisher's Note:** MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Copyright:** © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

Palm oil is a key source of edible vegetable oil extracted from fruits of the oil palm tree. It has emerged as an important feedstock and biofuel raw material [1,2]. Cultivated on millions of hectares in the world, oil palm has a noticeable part of the volume of world trade. The oil palm started its journey from West Africa and has become a hope for the economy of many countries. However, Southeast Asia is considered the hub of the palm oil industry, with Indonesia and Malaysia being the major exporters [3]. The growing demand for the palm oil threatens the future of the rain forests because expanding oil palm cultivations tend to replace existing forestry. To overcome the negative impacts of oil palm farming by making it a key element of building a future sustainable world, plant science faces three major challenges. The global average palm oil yield of 3.5 tons per hectare should be elevated to the full yield potential around 11–18 t. The tree architecture must be changed to low labour intensity and high mechanization of the harvest. Oil composition should be tailored to the evolving demands of oleochemical, fuel industries, and, most importantly, food [4]. Thus the use of technology is inevitable to deal with these challenges through the sustainable intensification of oil palm [5]. The application of powerful, intelligent ML methods has the potential to transform the current productive agriculture into sustainable agriculture. In precision agriculture, complex tasks such

as land classification, soil management, crop selection, seasonal variations, fertilizers management, crop yield forecasting, and yield gap analysis are effectively evolving with ML methods [6–9]. Because of incredible learning and high computational power, ML approaches have also been adopted in the field of oil palm for the mechanization of several tasks such as tree counting and plant health assessment [10,11]. However, the broader inspiration of modern agriculture is to examine all factors that cause gaps in crop yield. Agricultural productivity is subject to several stressors, including biotic and abiotic risk factors, many of which are intensified by a changing climate, thereby affecting durable sustainability. The productivity of tree crops such as oil palm is particularly complex. To comprehend and mitigate these risk factors, a collection of multi-layered large data sets is required. Additionally, advanced analytics is also critical to integrate those highly heterogeneous datasets to generate insights about the key constraints on the yields at tree and field scales. Since yield gap analysis deals with interdependent and irregular elements, it becomes difficult to map their dependencies with conventional methods. Thus, ML application is an efficient way to obtain precise outcomes and appropriate solutions for such convoluted problems. Previously, micro-components analysis was performed to study the overall yield gap by the commonly used “divide and conquer” ML strategy [12,13]. The studies concluded that the investigation of individual yield-reducing factors could provide insight into the overall impact. On the other hand, ML application to evaluate the correlation among different factors such as soil moisture and rainfall, day length and solar radiation, fertilization frequency, and plant growth provides better quantitative understandings [14–16]. The aforesaid statement necessitates bringing such research directions to light. In this regard, a domain-based scientifically structured review article can provide insightful information to new researchers based on compiled results. It enables depicting the state-of-the-art, highlighting the strategic matters raised by the scientific community and cluster analysis [17]. Researchers can identify research gaps with reference to discussed systems, concepts, and propositions. The review also helps to build a research roadmap with knowledge of formerly applied theories and existing techniques [18].

In the literature, some authors have reviewed the application of effective tools and ML techniques in oil palm. For instance, Chong et al. conducted a review on applications and practices of remote sensing for oil palm plantations monitoring [19]. Accordingly, the technologies developed for the processing of fruit and palm oil waste management by converting it to biofuel are reviewed in [20]. The application of ML for detecting nutrition deficits in oil palm with the help of proximal images are explored in [21]. Besides, the technical review of sensors and techniques for oil palm plants’ disease detection is performed in [22]. Subsequently, ML features are reviewed for automatic fruit grading with the help of image processing [23]. Recently, Rashid et al. reviewed ML application for yield prediction of different crops, including oil palm [24]. From the reviewed literature, it is observed that present studies did not systematically search the literature in most cases. Moreover, only specific aspects (such as yield prediction, crop monitoring, and nutrient deficits, etc.) of oil palm using ML are reviewed. Although the cited works provide deep insight into existing studies, to date, unidimensional reviews lack a full presentation of the overall ML-based research in the oil palm domain. The literature search reveals that a broader overview of the current research trends and the extent of ML application for oil palm is missing.

In need of a comprehensive analysis, this review article brings together the application of ML in the oil palm agricultural industry. To the best knowledge of the authors, there is no detailed review until now that has explored multiple aspects of the oil palm domain within the framework of artificial intelligence (AI). The present state-of-the-art review synthesizes all prominent studies published throughout an entire decade (2011–2020) concerning oil palm in the ML context. It stands out from other review studies as the literature search was performed through a systematic protocol to ensure unbiased data base retrieval. Moreover, the standard protocol was extended to perform qualitative and quantitative assessment of the searched literature. This technique allows detailed

and impartial insight into contemporary studies in the domain. The goal of writing this review was to analyse major progress, current trends, research gaps, and pioneering concepts intended for stimulating research on oil palm with the application of modern techniques. This research was envisaged to produce high-quality recommendations for novice researchers based on a scrupulous evaluation of the available works.

The remaining article is organized as follows. Section 2 provides the background of oil palm and ML. Section 3 explains the review scope and the proposed methodology. It provides the step-by-step procedure followed during the design and implementation phase. Accordingly, results are demonstrated in Section 4, while detailed discussion is presented in Section 5. Finally, Section 6 concludes the article.

## 2. Background of Oil Palm and Machine Learning

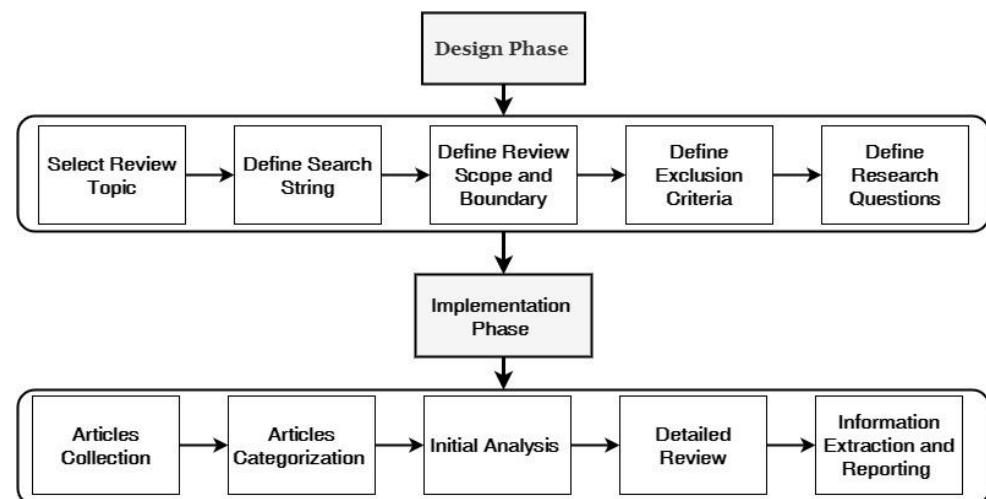
The *Elaeis guineensis*, commonly known as oil palm, is a single-stemmed branchless tree, which takes several years of investment and labour work before producing harvestable oil containing fresh fruit bunch (FFB) [25]. Two different kinds of vegetable oils are produced from oil palm fruit at first, namely, crude palm oil (CPO) and palm kernel oil. CPO is extracted from the pulp (mesocarp) of fruit, while its seeds provide palm kernel oil. Average annual raw crop production remains 12–18 tons  $\text{h}^{-1}$  in industrial oil palm under favourable conditions [26]. Besides its profitability and high production, some complex characteristics distinguish it from other crops. Unlike annual, biennial, and perennial crops, the oil palms are permanent plants [27] harvested twice a month throughout the plant's lifespan (except the initial growing period) [28,29]. Although some seasonal variations exist in oil palm yield, this divergence is limited to production capacity [30]. The seasonal impacts are considered a noticeable factor in contributing to yield decline. Some additional yield-reducing factors are fertilizers' limitations, irrigation limitations, pests, and infections. Careful regular analysis for disease assessment, fertilization, and timely harvesting contributes to high oil palm yield. Conversely, inappropriate field management and incorrect harvesting strategies restrict the oil quality and quantity [31]. Besides all efforts to enhance oil palm crop production, the yield-reducing factors influence the outcomes significantly. However, the factors mentioned above are variable and interdependent but are not unpredictable. Indeed, these sensitive issues are associated with some complicated parameters. As for fertilizer, its type, adequate quantity, and application frequency require precise optimization. At the same time, the estimation of harvesting time [32] is achieved with predictive analysis. Similarly, during disease assessment, healthy and unhealthy plants are classified. A decisive solution to such a complex problem is rendered in machine learning algorithms (MLAs) that perform prediction, classification, clustering, and optimization. Depending on research requirements, an ML model can be predictive or/and descriptive. A predictive model is used to forecast what is likely to happen in the future. On the other hand, a descriptive model gathers knowledge from the collected data to describe what happened in the past [33]. For solving the problems at hand, selection of the right algorithms is the key. Moreover, the utilized tools and chosen algorithms need to have the capability of handling bulk data. Aided by such intelligent algorithms, well-calculated field management strategies can boost the revenue through enhanced yield and optimized expenses. In addition, the oil palm as a profitable crop also needs technology-oriented tools (to reduce labour costs) and cerebral systems (to avoid the risk of human error) for decision-making.

## 3. Review Scope and Methodology

This work reviews the existing studies that apply different ML techniques to serve the oil palm agricultural industry from a problem-solving perspective. The proposed ML methods are explored while emphasizing recent research trends. We do not intend to criticize previous contributions; moreover, it is not a comparative analysis to support any specific research work or algorithm. To the best knowledge of the authors, barely any attempt could be discovered in the existing literature that reviews "application of MLAs in

the field of oil palm” as a special case because research on oil palm is limited. Such a dearth can be acknowledged as a relatively novel, yet promising, application field. Moreover, this review only considers MLAs that offer enhanced estimates rather than optimization algorithms taking a long time for finding global optima. In addition, it does not consider supplementary research on the advantages/disadvantages of oil palm consumption as well as field management strategies that sustain without ML. Those topics are not in the scope and do not fit the level of its specified target.

For this review, the methodology was based on a systematic literature review approach to perform a hybrid qualitative–quantitative analysis. The motivation for using the consolidated analysis was to provide a better understanding of the results by weighing the pros and cons of both techniques [34,35]. This method is susceptible to minimize bias, enhance transparency, and preserve flexibility [35,36]. In this study, we organized our empirical outcomes and answers to the “research questions” that are a central part of this research protocol. Moreover, a dual phased systematic method is proposed where each phase comprises five major sections as sub-steps. The procedure is described in Figure 1.



**Figure 1.** Dual-phase systematic review procedure. The procedures followed at the first stage are part of the design phase.

### 3.1. Design Phase

#### 3.1.1. Review Topic Selection

To conduct this review, first of all, a domain-specific less intact research topic was carefully chosen after a deep study of existing literature related to latent advancement in oil palm with the help of ML. In the quest of exploring the studies that have been published in the domain of oil palm and ML, this review topic is appropriate to analyse several dimensions.

#### 3.1.2. Defining Search String

The basic searching was done by an automated search. The starting input for the search was “oil palm” AND “machine learning.” Articles were retrieved, and abstracts were read to find the synonyms of the keywords. The search input was used to obtain a broad view of the studies. From the basic search experiment, a more complex search string was built in order to avoid missing relevant studies. The final search string was as follows: (“machine AND learning” OR “deep AND learning” OR “artificial AND intelligence”) AND (“oil AND palm” OR “Elaeis AND guineensis”). After executing iterated combinations of defined search strings in five databases, 1060 studies were retrieved through advance search with the title, abstract, and keywords. The “anywhere” option in Google Scholar and Springer Link was selected. Inserted keywords included “oil palm,” “Elaeis guineensis,” “machine learning,” “deep learning,” and “artificial intelligence” as search strings according

to the procedures defined in [37] with the AND keyword for the exact combination of two strings and OR for flexible searches. The initial search was performed based on the title, abstract, and keywords; however, the full text was considered for final selection, categorization, and information extraction.

### 3.1.3. Defining Review Scope and Boundary

The available literature on the topic was collected within a unique range of ten years. Only technical papers that propose the application of any MLA to explore oil palm were considered. On the other hand, all the articles that did not apply ML for agricultural oil palm as a special case were discarded. This was done to narrow down the search because applications of ML occur in numerous fields that do not serve our domain-specific investigation. For instance, a study was not included if it investigated public opinion (consumer perception) on the impacts of palm oil using ML methods [38].

### 3.1.4. Defining Exclusion Criteria

To incorporate only relevant publications, several exclusion criteria were defined to set boundaries for the review. In the systematic procedure of article selection, seven exclusion principles were followed to filter the database as described below.

- Exclusion Principle 1—publication that is not related to the agricultural/industrial sector of oil palm with ML application.
- Exclusion Principle 2—publication is written in any language other than English.
- Exclusion Principle 3—publication is already retrieved/is a duplicate.
- Exclusion Principle 4—full text/abstract of the publication is unavailable.
- Exclusion Principle 5—publications that are survey/review articles reviewing old publications.
- Exclusion Principle 6—work that is published before 2011.
- Exclusion Principle 7—publication is published after 2020.

### 3.1.5. Research Questions

The review objective was based on seven research questions (Qs).

- Q1—what is the annual rate of publications that apply machine learning to oil palm?
- Q2—what are research trends in this domain?
- Q3—which MLAs have been highly used during the previous ten years to serve oil palm?
- Q4—which countries are actively involved in research on the application of machine learning in oil palm?
- Q5—what datasets/sources are used to investigate oil palm with machine learning?
- Q6—what are the key input features preferred by researchers?
- Q7—how are applied algorithms/ models evaluated to guarantee the significance and rationality of the outcomes?

## 3.2. Implementation Phase

This section covers the procedures to acquire relevant literature. Primarily, the articles that contained standalone or hybrid ML models integrated with other methods, precisely applied on oil palm, were searched. The choice of articles included literature with macro-level (country or state), meso-level (entire plantation), and micro-level (tree or part of tree) oil palm assessment with multidimensional application of statistical, ML, or deep learning (DL) algorithms.

### 3.2.1. Article Collection

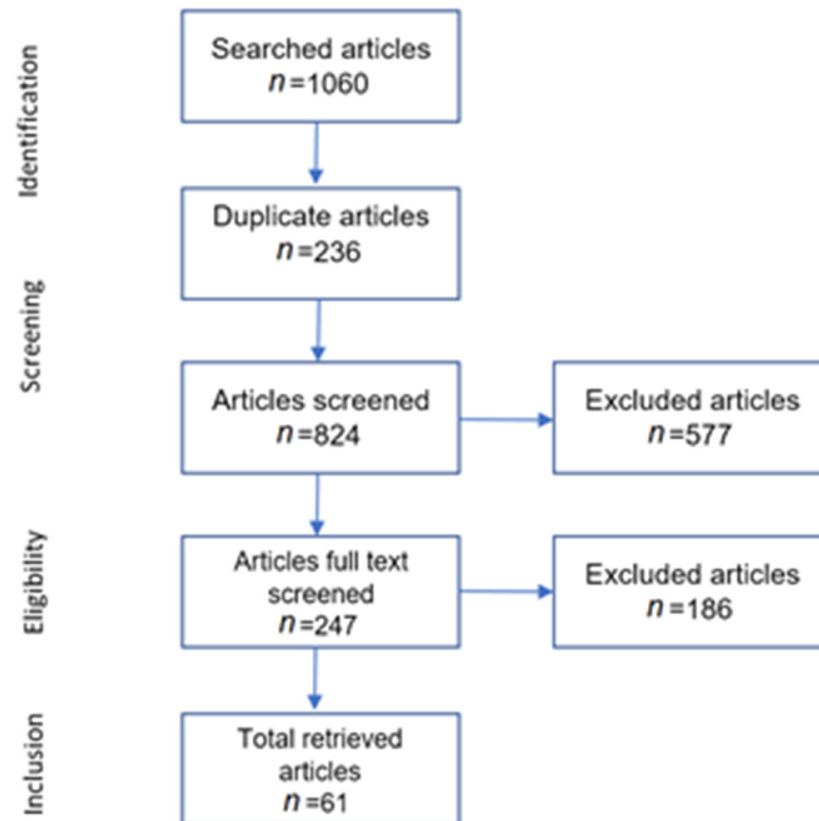
The article collection phase was inaugurated after streaming through the entire set of exclusion criteria. The candidate publications were identified and stored in a distinct database. Initially, we retrieved articles from different sources, including Scopus, Science



Direct, Springer Link, Web of Science, and Google Scholar. To limit the literature search, a date filter was applied to extract articles within the defined period.

### 3.2.2. Initial Analysis

From the available database, the most significant and relevant literature was extracted after full-text reading to finalize the candidate list. Consequently, the number of remaining publications was reduced to 61, and we decided to analyse all the publications further regardless of the impact factor and the number of citations. Figure 2 describes the adopted PRISMA protocol [39].



**Figure 2.** Diagram of PRISMA flow.

### 3.2.3. Articles Categorization

The articles categorization phase divides the entire database into two main groups: classification and regression. The grouping is performed based on analysis type instead of technical classification. Both groups are further divided into different subcategories considering the prime research objective(s) according to the strategy stated in [40]. It should be noted that some works containing multiple research objectives are included in more than one class; however, they were counted only once to avoid miscalculations.

### 3.2.4. Detailed Review

This step includes a survey of literature to be reviewed for answering the research questions. All articles were deeply examined on an individual level for information extraction as well as to obtain unified statistics of the entire database.

### 3.2.5. Information Extraction and Reporting

Referring to Q1, the annual distribution of extracted literature was examined. Accordingly, the scope and objectives of the study were considered regarding Q2. Subsequently, a detailed analysis of the methodology stated in the reviewed articles was performed to

extract statistics of applied MLAs vis-à-vis Q3. Afterwards, the affiliations of authors and all co-authors responded to the fourth research question (Q4). Further study of materials and methods along with results was performed in order to discover the input data, key input features, and performance evaluation techniques to comply with Q5, Q6, and Q7, respectively.

#### 4. Results

This section presents the outcomes of the proposed methodology described in the preceding section. The findings portray yearly trends of publishing articles in the discussed domain, the main problems and researchers' objectives, the adopted algorithms, the geographical information, the data sources, and the model performance evaluation methods. Regarding the annual distribution of published work, trends indicate that a higher number of research articles were published during the years 2014, 2016, 2018 and 2019, while, in the year 2015, the number of articles dropped too low. The yearly distribution of research publications is demonstrated in Figure 3. This outcome responds to the first research question (Q1).

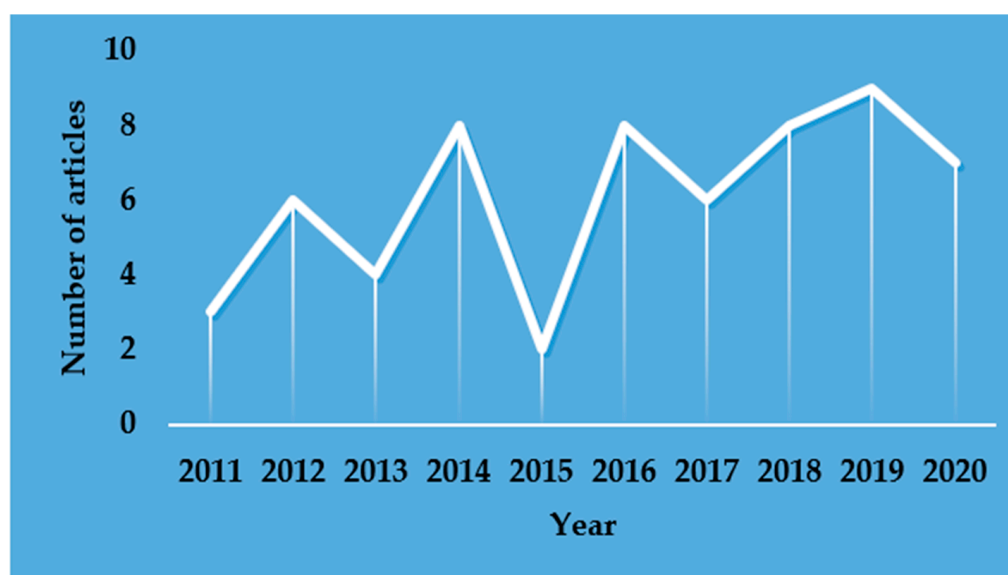


Figure 3. Annual distribution of reviewed articles.

The outcomes addressing Q2 provide further insight into the current research trends in the application domain. The results suggest that approximately 84% of work in the literature was dedicated to classification techniques, while the remaining 16% of work was focused on regression analysis for forecasting. The important research objectives in the classification category were tree/land detection, fruit classification, disease detection, multipurpose classification, and canopy monitoring with unmanned aerial vehicles (UAV). Among the preceding subcategories, one class, namely, multipurpose classification contained some articles with diverse (miscellaneous) objectives. On the other hand, various objectives performed under regression analysis included the prediction of harvest time, fruit yield, oil yield, palm oil prices, and seasonal impacts on production. The main categories and the subcategories with approximate proportions are presented in Figure 4.

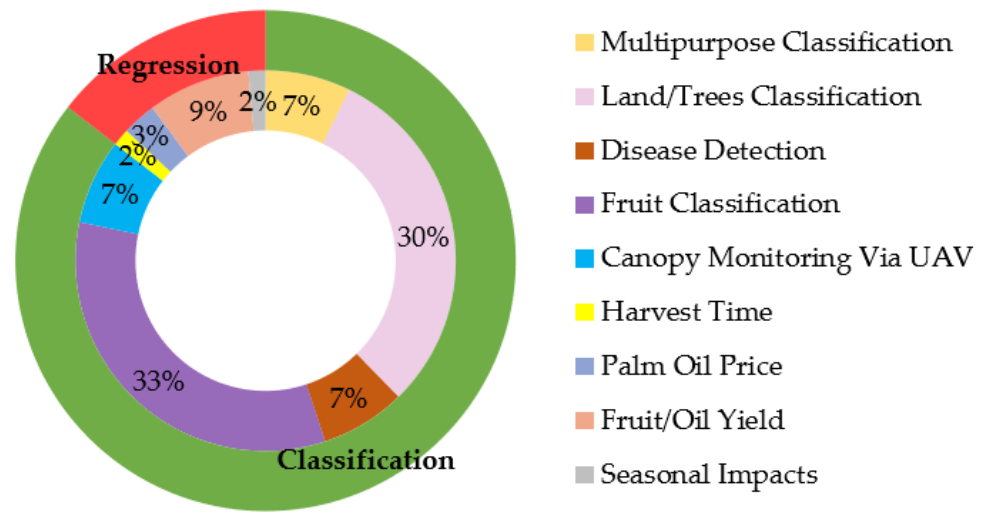


Figure 4. Articles categories and objective-based distribution.

Further assessment of the reviewed database was performed to find out highly used MLAs (referring to Q3). In the overall literature, support vector machine (SVM) was the highest applied algorithm followed by artificial neural network (ANN), random forest (RF), regression, classification and regression tree (CART), and convolutional neural network (CNN), respectively. The most popular MLAs and their usage frequency in question under discussion are presented in Figure 5.

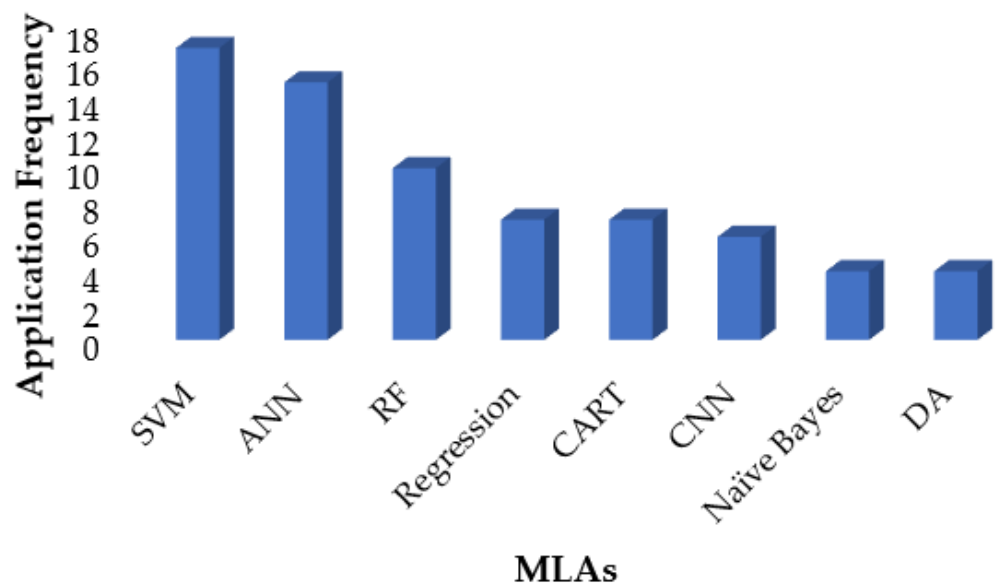


Figure 5. Highly adopted MLAs.

To address Q4, we extracted the authors’ affiliations to achieve the geographical distribution of the reviewed publications. All authors and co-authors were assumed to represent the country mentioned in the author’s profile. The contribution of researchers from all countries as per the number of publications is presented in Figure 6.



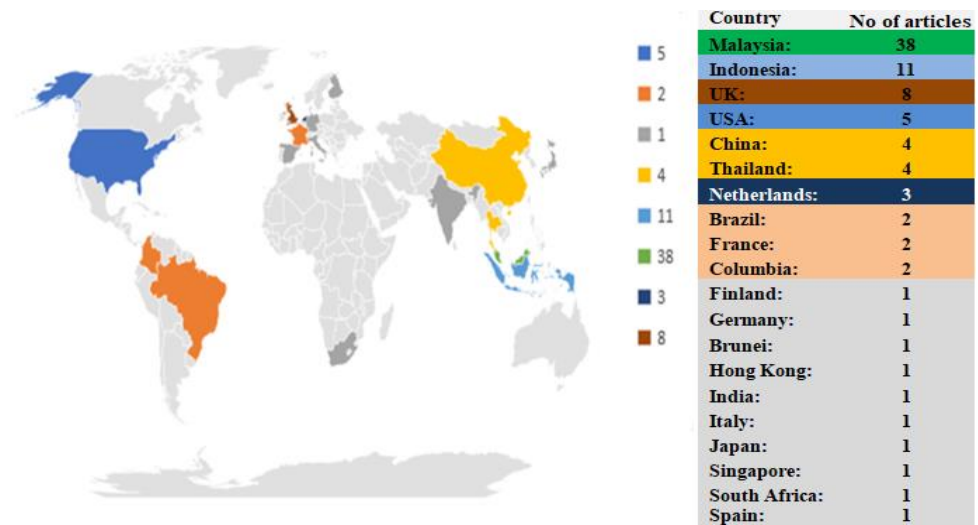


Figure 6. Geographical distribution of reviewed articles.

To address the fifth research question (Q5), articles were dug to explore the data sources and utilized data sets contingent upon explicitly stated data sources. This investigation was in support of researchers and practitioners to obtain information on data availability because the performance of ML models highly depends on the quality and appropriate quantity of the input data. It was discovered that multiple sources providing satellite images were mentioned and that field-specific self-collected data is widely used. Figure 7 demonstrates the approximate proportion of sources and data sets in the reviewed literature. The miscellaneous category contains multiple datasets for the range of parameters that were not distributable among other subcategories.

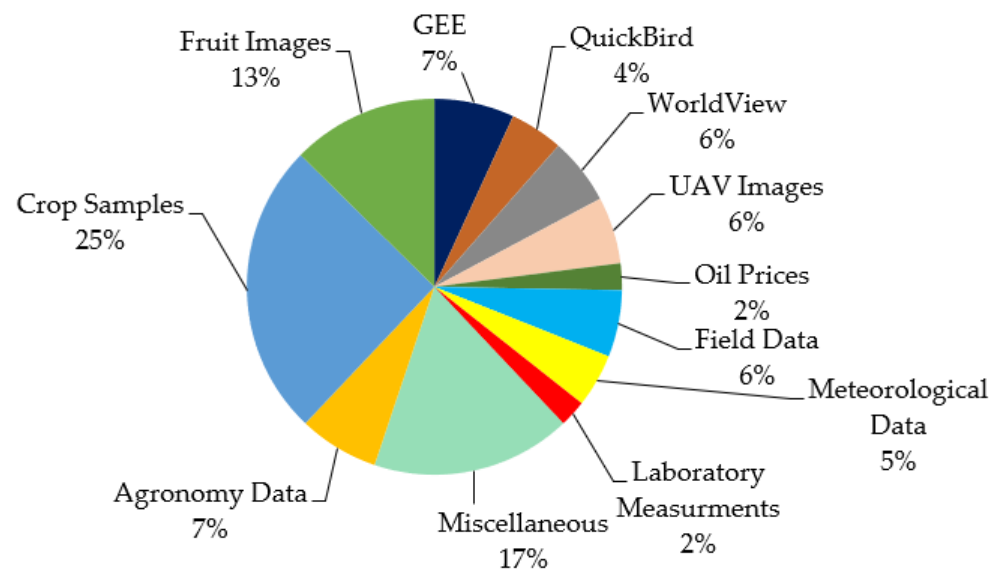


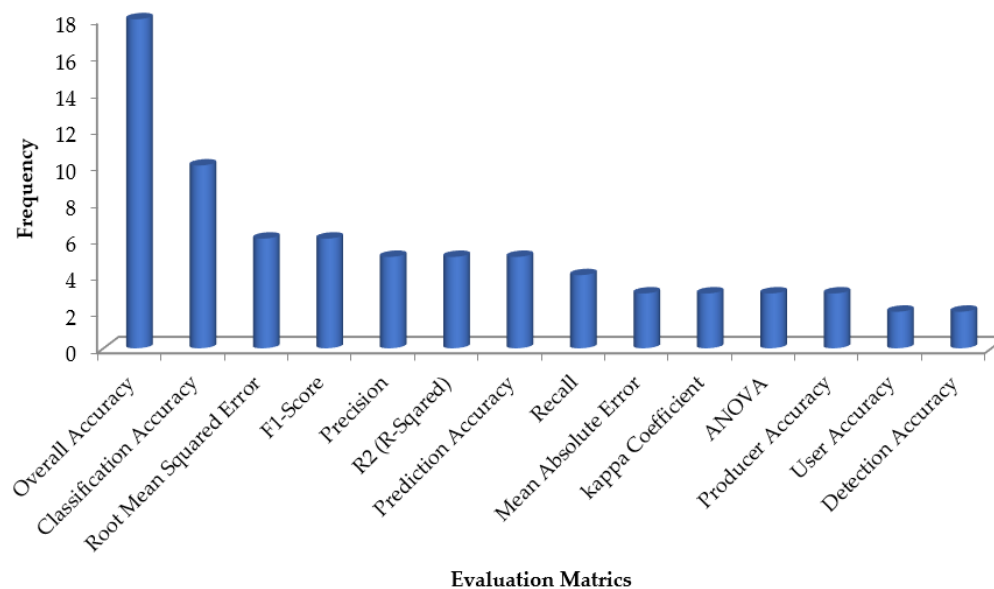
Figure 7. Datasets/sources mentioned in the articles.

The response to Q6 describes the input features used by the researchers. The information is beneficial for feature extraction to develop effective and efficient models. We combined all input variables against each relevant category. For better understanding, the results are tabularized, as can be seen in Table 1.

**Table 1.** Input features.

Papers Category	Input Features
Multipurpose classification	Tree crowns and categorical features; thermal images; quantitative features; climatological features; ratios of kernel to fruit; shell to bunch; shell to fruit; fruit to bunch; messocarp to fruit; oil to dry messocarp, and oil per palm; chlorophyll-sensitive wavelengths; FFB colours; electrical properties of leaves; and levels of potassium (K), nitrogen (N), phosphorus (P), magnesium (Mg), and calcium (Ca)
Disease detection	Tree crown size and categorical features, spectral reflectance, and leaves and stem colours
Prediction/estimation	Fruit age, plant life, age factor, normalized difference vegetation index (NDVI) average value, daily CPO prices, monthly closing prices of oils, oil palm production, benefits, plant scale index, sex determination, inflorescence abortion, foliar nutrient composition, FFB yield, growth, respiration, and meteorological variables
Land cover/tree detection	Positive/negative histogram of oriented gradients (HOG), crown size, images of oil palm, built-up, bare land, water, forest
FFB Analysis	Flavonoid, anthocyanin content, fruit colour, fruit size, hue, saturation, intensity, contour lines, blue-to-red fluorescence ratio (BRR_FRF)
UAV for canopy monitoring	Tree crown, stems length, tree colours, crown size, bare land, water

Q7 was addressed by analysing models’ performance evaluation parameters statistics. The ranking of evaluation parameters based on usage frequency in the reviewed literature is shown in Figure 8.



**Figure 8.** Most frequent evaluation parameters.

All above results provided insights into statistics of the literature on ML and oil palm to encounter our seven defined research questions. The specified qualitative review of individual articles from the subcategories can be extracted from Tables 2–7, and the brief precedent description of important studies is provided prior to each table.

#### 4.1. Multipurpose Classification

Some articles were added to multipurpose classification due to unique objectives. Articles in this category are the works that have conducted the technical research predominantly focused on classifier's performance. One of the studies included in this group explains the results of various marker systems and modelling procedures for the implementation of genome selection (GS) in the introgressive hybridization of the dura family. The discussed dura family "Deli x Nigerian," the resultant of a "Deli dura x Nigerian dura," was considered in the analysis [41]. The observed family is a valuable source of new palm trees with higher oil yields and improved bunch characteristics. The work in [42] proposed Jenks natural breaks (JNB) for the classification of chlorophyll sufficiency levels and relative chlorophyll content. The best subset of frond number, chlorophyll-sensitive wavelengths, and the classifier to categorize the chlorophylls according to the nominated sufficiency levels were suggested using a hyperspectral remote sensing platform. Correspondingly, the use of hyperspectral sensing combined with imbalance approaches and MLAs to track the nutrients levels of mature oil palm is highlighted in [43]. Hyperspectral spectroscopy has emerged as a promising alternative to conventional foliar analysis in assessing the nutritional status of oil palms, as the former one is costly and time-consuming. Details about classification techniques used in the oil palm field are provided in five different tables. Table 2 contains literature on multipurpose classification techniques applied to examine oil palm.

**Table 2.** Multipurpose classification.

Articles	Dataset	Model/Algorithm(s)	Objective
[44]	Thermal and meteorological dataset, labeled images of area of interest (AOI)	RF, K nearest neighbour (KNN), SVM	Classification of female inflorescences' anthesis stages in oil palm
[41]	Average phenotype	Ridge regression-best linear unbiased prediction (RR_BLUP), Bayes A, Bayes B, $C\pi$ , SVM, ridge regression (RR), RF, least absolute shrinkage and selection operator (LASSO)	Evaluation of marker systems and methods for genomic selection of oil palm
[42]	Tree images, leaf spectral reflectance, and fertilizers' application experiments	Decision tree (DT), RF	Calculation of chlorophyll sufficiency levels in mature palm using hyperspectral remote sensing
[43]	Leaf samples, nutrients	Logistic model tree (LMT), naïve Bayes (NB), synthetic minority over-sampling technique (SMOTE), adaptive boosting (AdaBoost)	Oil palm's macronutrients classification
[45]	FFB samples	CART	Fruit ripeness identification
[46]	Laboratory measurements	ANN, linear model (LM), gradient descent algorithm (GDA)	Modelling of dielectric properties of oil palm fruitlets

#### 4.2. Disease Detection

Diseases in oil palm are a major threat to global food supply and security. Timely detection of infected trees is a difficult task due to the lack of appropriate monitoring

resources for plants health assessment. Basal stem rot (BSR) is a common plant disease instigated by the *Ganoderma boninense* (*G. boninense*) fungus, which passes on the infection to oil palm plantations and causes a significant economic loss. In need of novel disease detection techniques in order to reduce oil palm losses caused by BSR, images along with field observations are utilized to perform classifications of infected oil palm estate through different ML algorithms in [47]. The purpose of this study was to use WorldView-3 imagery to predict disease severity using supervised learning algorithms, as well as to identify the typical indicators of BSR disease in oil palm at various levels of infection severity. Infection levels are classified from mild to severe using three different classes with the help of various classification algorithms. Similarly, Khaled et al. considered the impacts of BSR by classifying the healthy and unhealthy leaves from collected samples [48]. The feasibility of using electrical properties, namely, impedance, capacitance, dielectric constant, and dissipation factor, for rapid detection of the BSR disease in oil palm trees was explored. The detailed analysis was performed using five different MLAs, of which three models were exploited for features selection and two performed classification. Nevertheless, early identification of *G. boninense* infection is the technique for controlling BSR because contemporary methods do not guarantee complete recovery after severe infection. The key goal of a different study was to examine how well ML models can detect BSR infection in oil-palm-cultivated areas. The BSR disease distribution map was also created through MLA combined with remote-sensing techniques [11]. The rotten bunch disease was detected in [49]. Another study found that healthy trees can be differentiated from BSR-infected trees through intelligent algorithms even in early-stage with less visible symptoms [50]. Table 3 includes itemized material from specific articles that performed classification for the purpose of disease detection in oil palm plantations.

**Table 3.** Disease detection.

Articles	Dataset	Model/Algorithm(s)	Objective(s)
[11]	Quickbird imagery	SVM, RF, CART	Modelling of basal stem rot disease
[50]	Weather data, infected trees test, spectral reflectance data, detached leaflets from fronds	ANN	Early detection of ganoderma basal stem rot
[48]	Impedance, capacitance, dielectric constant, and dissipation factor in infected trees	SVM, RF, and genetic algorithm (GA) (for features selection), ANN, SVM for classification	Spectral features selection and classification of infected oil palm leaves from BSR disease
[49]	Disease, symptoms, treatment	NB	The diagnosis of disease in oil palm
[47]	Worldview-3 imagery, real time field imagery	DT, RF, SVM	Severity of BSR in oil palm farms

#### 4.3. UAV for Canopy Monitoring

The concept of UAV for crop monitoring is not quite new in modern agricultural practices. Aerial colour and colour-infrared imaging have been used to track crop growth for more than 50 years. These methods are currently being re-evaluated for analysis in precision agriculture. UAVs are becoming more popular due to their low cost and ability to fly on low altitudes, which increases the spatial resolution [51]. The aerial imagery can be acquired rapidly during crucial periods of crop development. For oil palm field monitoring through UAV and classification of crop segments based on resulting images, linear regression was performed in [52]. Similarly, ref. [53–56] collected images of oil palm from UAV for counting or detection of oil palm trees by applying different ML techniques.

The articles that suggest automated oil palm canopy monitoring using UAV are included in Table 4 as an imperative part of this review.

**Table 4.** Automated canopy segmentation/crop monitoring using UAV.

Articles	Dataset	Model/Algorithm(s)	Objective(s)
[52]	Images of 2-, 4-, and 7-year-old trees	Linear regression (LR)	Automatic canopy segmentation
[53]	Plantation images	Visual geometry group-single shot detector (VGG-SSD), faster-RCNN, YOLO-V3, Retina-net, Mobilenet-SSD	Fast and robust detection of oil palm
[54]	Oil palm images	Histogram of oriented gradients (HOG-SVM), SVM	Detecting Individual oil palm tree
[55]	Oil palm images	SVM	Counting oil palm inventory
[56]	Oil palm images	CNN	Oil palm tree detection

#### 4.4. Prediction/Estimation

There is a plentiful scope for research on forecasting methods to support in many phases during oil palm cultivation's lifecycle. The studies stated that the current issue in agricultural oil palm is the high difference between potential and actual yields. However, the broad application of multidisciplinary ML methods can overcome this issue. Many factors influence crop yield, including crop genotype, environment, and management strategies. Year-to-year and location-to-location differences in crop yield are greatly influenced by changing environments, both spatially and temporally. Accurate harvesting time, seeds' sowing time, irrigation requirements, price prediction, and yield estimation is extremely beneficial to the production of global food under such circumstances. Decisions on appropriate import and export can be made on the basis of reliable forecasts [57]. In order to implement an appropriate harvesting strategy, oil palm harvest time is predicted based on fruit growth and ripeness level with regression analysis [58]. An oil palm prediction model to estimate production from cultivated area images and tree age estimation is proposed in [59]. Forthcoming CPO prices are predicted in [60,61] based on historical trends in CPO prices. The performances of two regression models trained from historical data of oil palm production are compared to predict future production in [62]. At the same time, a simple prediction model has been proposed to simulate the impacts of the environment on variance in the inflorescences and the number of harvested bunches [63]. Similarly, another study introduced a standard simulation model to predict the growth of oil palm trees and the potential yield [64]. The impacts of foliar nutrients' compositions have been analysed to model and predict the oil palm yield of Malaysia, inclusive of historical trends in FFB yield [65]. Likewise, annual oil palm yield is predicted, and climate impacts are explored, in [66]. Finally, small-scale oil palm yield has been predicted with the help of historical yield data and multiple environmental factors [67]. The above-discussed literature is part of Table 5.

**Table 5.** Prediction/estimation.

Articles	Dataset	Model/Algorithm(s)	Objective(s)
[58]	Field aata, FFB samples	Discriminant analysis (DA), polynomial regression (PR)	Identifying ripeness and forecasting harvest time of oil palm
[59]	SPOT6 imagery	ANN, linear regression	Estimation of the amount of oil palm production
[60]	CPO historical prices	ANN, autoregressive fractional integral moving average (ARFIMA), adaptive neuro-fuzzy inference system (ANFIS)	Forecasting on crude palm oil prices
[61]	CPO historical prices	SVR, holt winter exponential smoothing	Multivariate time series forecasting of crude palm oil price
[62]	Time series data	Support vector regression (SVR), ANN	Prediction of oil palm production
[63]	Crop characteristics, climatic data	ECOPALM	Prediction of seasonal variations in FFB production
[65]	FFB yield—the foliar nutrient composition	ANN	Modelling Malaysian oil palm yield
[64]	Description of water, solar radiation, and nutrients	PALMSIM	Simulating potential growth and yield of oil palm
[67]	Soil fertility, water, weather, and historical FFB yield	Bayesian network	Predicting future FFB yield
[68]	Region of interest (ROI) images, rule-based expert opinion	Rule-based expert system (RBES), KNN, SVM, ANN	Oil palm FFB ripeness prediction

#### 4.5. Land Cover/Tree Detection

Oil palm is among the fastest-growing crops in terms of agricultural land use, and their development has been linked to substantial damage to the environment. As a result, this crop often appears in open and procedural discussions that are hindered or skewed by a lack of reliable conservational data. The shortage of consistent cultivation information in particular has continued to be a source of concern. Recent advances in remote-sensing data access and the ability of ML have played a remarkable role to address this issue. In this regard, the study in [10] has obviously contributed not only to oil palm tree detection, but it also incorporates ML for tree counting with adequate accuracy. Tree counting along with young and mature tree identification is carried out by Okoro et al. [69]. Some other studies [70–73] also performed similar work related to tree counting, age estimation, or tree detection but differ in study level, data sets, and algorithms. Another wide variety of researchers seems more concerned about oil palm area mapping, globally and locally, as [74,75] and many more have proposed models to monitor oil palm area expansion, covered land, plantation detection, patterns of land use for oil palm plantation, above-ground biomass, oil palm fields observation, and so on. Particulars on the above-discussed contributions are provided in Table 6. It covers all reviewed articles implementing various models to detect land cover, oil palm cultivation, or expansion. However, no model has been proposed to differentiate oil palm from other similar trees like date palm and coconut



palm. This research gap may cause inaccurate classification of satellite images in case of analysing diversified cultivation on a large scale [76].

**Table 6.** Land cover/tree detection.

Articles	Dataset	Model/ Algorithm(s)	Objective(s)
[74]	GEE Sentinel-1,2 images	DEEPLABV3+ CNN	Worldwide map of oil palm plantations
[75]	GEE images	RF	Recording the spatial allocation of oil palm
[77]	WorldView-3	CNN	Detecting young and mature oil palm trees
[69]	GEE images	SVM	Monitoring oil palm cultivation monitoring
[78]	Images captured at site	Viola and Jones detector	Oil palm map
[79]	Calculated biomass values, forest type, and soil information	DA, logarithmic regressions	Estimating aboveground biomass in oil palm plantation
[80]	Satellite imagery, time series data from MODIS, field data	Moving average (MA), DT	Detecting land cover conversion to oil palm
[10]	Remote sensor images	CNN	Tree detection and counting
[81]	Images from Kaggle	DenseNet, DenseNet with saliency and semantic parsing (SSP)	Industrial oil palm monitoring
[82]	Earth Explore images	SVM	Oil palm distribution mapping
[70]	QuickBird satellite images	Faster-CNN	Detecting oil palm trees
[73]	QuickBird satellite images	TS-CNN	Detecting oil palm trees
[71]	WorldView-3, LiDAR	SVM, RF	Oil palm tree counting and age estimation
[83]	GEE Images, shuttle radar topographic mission (SRTM), NDVI, normalized difference water index (NDWI), digital elevation models (DEM)	SVM, CART, RF	Oil palm mapping
[72]	QuickBird imagery	Vegetation indices, semi-variogram computation	Detection of oil palm trees
[84]	GEE images	SVM, RF, CART	Monitoring oil palm farms in Malaysia
[85]	Land cover map, WorldView-2 image, field data	SVM and maximum likelihood classifier (MLC)	Mapping of oil palm
[86]	Site-specific agrometeorological data	Dempster-Shafer Inference	Irrigation management in oil palm crops
[87]	Palm trees images	Logistic regression	Validation of an oil palm detection system

#### 4.6. Fresh Fruit Bunch Analysis

Prior analysis of oil palm's FFB based on its ripeness stage and other characteristics is needed for sorting and harvesting to obtain the supplementary quantity and upheld oil quality that is compulsory to meet international standards. Ordinarily, oil palm harvest time is defined by its fruit maturity stage that guarantees a high oil extraction ratio (OER) by avoiding over-ripe, under-ripe, or unripe fruit harvest. As FFB is the first raw product gathered from oil palm cultivation, its sensitivity towards the ripeness stage motivated too many researchers for utilizing machine vision to determine fruit age, size, quality, etc. that is vital to ripeness stage identification. A large portion of the reviewed literature consists of methods and models for ripeness detection. The study in [45] makes a significant contribution by demonstrating that the blue to red fluorescence ratio-index can be used to evaluate the oil palm FFB maturity. Real FFB samples were observed for differentiating among maturity stages. On the other hand, ref. [88] processed FFB images to perform classification for fruit maturity level identification. Using the artificial fuzzy logic and red-green-blue (RGB) colour model, an automatic fruit grading system for oil palm was proposed in [89]. The resulting grading system can differentiate among various types of oil palm fruits based on colour intensity to describe ripeness level. Another study used a hyperspectral-based method for ripeness detection of FFB [90]. Conversely, Makky et al. proposed a multiclass classification method integrated with machine vision to develop a fruit grading machine [91]. By applying mathematical equations and regression analysis, ref. [92] describes a method to estimate the maturity stage and age of oil palm FFB based on its location in the phyllotaxis as an added element to confirm the maturity level. Another straightforward model was introduced to build a tool for oil palm fruit ripeness detection in [93]. The method of image segmentation to discriminate the region of oil palm fruit was proposed in [94]. To accomplish the segmentation method, Gaussian filtering has been developed for reducing image background noise. Likewise, an edge detection process has been introduced to obtain an outline of the oil palm fruit. Throughout [45,46,88–92,95–107], it was discovered that although adopted methods, ideas, and input data varies, key objectives largely remained on the analysis of oil palm fruit, which can be seen in Table 7.

**Table 7.** Oil palm fruit analysis.

Articles	Dataset	Model/Algorithm(s)	Objective(s)
[93]	FFB images	LR, ANN, principal component analysis (PCA)	Oil palm fruit grading
[95]	FFB images	ANN	FFB ripeness detection
[96]	FFB images	SVM, NB	FFB ripeness grading
[89]	Fuzzy logic	Rule-based classification	Oil palm fruit grading system
[94]	FFB images	Background segmentation, <i>Gaussian filtering</i>	Segmentation of oil palm FFB images
[97]	FFB images	DA, ANN, PCA	Features extraction of oil palm FFB
[46]	Laboratory measurements	ANN, LM, Gaussian discriminant analysis (GDA)	Modelling of dielectric properties of fruitlets
[98]	FFB, flavonoïdes, anthocyanines	CART, ANN, stochastic gradient boosting trees (SGBT)	Determining FFB ripeness

Table 7. Cont.

Articles	Dataset	Model/Algorithm(s)	Objective(s)
[45]	FFB samples	CART	Fruit ripeness identification
[99]	FFB images	PCA- ANN	FFB ripeness determination
[100]	FFB sample	Oil palm fruit sensor concept	Grading of FFB
[101]	Palm fruitlets	Multi ANFIS	Inference system for dielectric properties of FFB
[45]	FFB samples	CART	FFB ripeness detection
[88]	FFB images	ANN	FFB ripeness detection
[102]	FFB images	Back propagation, learning vector quantization	FFB ripeness determination
[90]	FFB images	ANN	Fruit grading
[103]	FFB images	K mean clustering	FFB growth determination system
[104]	Field data	Unified modeling language (UML diagram)	Fruit grading
[105]	FFB images	SVM, ANN, ALEXNET	Fruit ripeness grading
[106]	FFB images	ANN, KNN, SVM	FFB ripeness grading
[92]	FFB acceptance rules, images	Mathematical regression	Grading machine for oil palm FFB
[107]	FFB age, maturity, spiral leaf images	Mathematical equations	FFB maturity stages determination
[68]	Region of interest (ROI) images, rule-based expert opinion	Rule-based expert system (RBES), KNN, SVM, ANN	FFB ripeness determination

## 5. Discussion

### 5.1. Results-Based Discussion

This subsection provides the discussion of obtained results referred to in the preceding section. As observed, the annual number of published articles is progressive but lacks variety in objectives. Around 33% of the research focused on fruit classification methods. Mainly, the authors performed FFB analysis for fruit ripeness detection, which is crucial to estimate harvesting time. In most of these studies, the images of FFB and real oil palm fruit samples were collected to detect differences among four ripeness stages, namely unripe, under-ripe, ripe, and overripe fruit. However, due to some common features between consecutive stages, ripeness identification becomes difficult with conventional classification approaches. Some authors also proposed useful features for image processing and oil palm fruit classification. In the category of work carried for health assessment, namely, the disease detection group, several MLAs were applied. Although 7% of the research revolved around the detection of disease and infection in oil palm, the amount of work done was insufficient considering the importance of the problem. Similarly, another 7% of research was dedicated to canopy monitoring using UAV, while the multipurpose classification group with various assorted objectives also constituted 7% of the total number of reviewed articles.

Another noticeable trend according to objectives in reviewed works was land and trees observation, which contributed to 30% of the reviewed literature. The impetus to use

ML for this purpose was derived from two dominant needs. First, the difficulty in counting oil palm trees manually and second is to monitor the expansion of oil palm cultivation as it is considered a threat to tropical forests, biodiversity, and associated ecosystems. For better supervision, the reviewed studies used classification techniques with remote sensing and machine vision to monitor oil-palm-cultivated areas. Remote sensing integrated with classification models is widely applied for tree detection, tree counting, oil palm tree identification, and so on. Some authors preferred multiclass classification [108] techniques over binary classification [109] for oil palm detection.

Apart from classification, regression analysis is also another prevailing research area. This category is mainly concerned with applying ML for estimation and prediction. In the reviewed studies, various regression models were proposed to achieve different objectives of forecasting such as oil palm growth, oil quantity, FFB yield, palm oil prices, and seasonal variations. Although progress in forecasting with ML is praiseworthy, it is not satisfactory. The results of the review suggest that forecasting with ML in oil palm is still in its early stages, and several obstacles have to be overcome. Obviously, ML is underutilized in the oil palm industry for descriptive and predictive analysis. Nonetheless, it is anticipated that, sooner or later, ML is likely to contribute to solving more challenging problems in agriculture and specifically for the field of oil palm as part of agriculture 5.0 [110].

Regarding applied algorithms, deep insight into the results shows that the application of algorithms differs in relation to the objectives in defined categories. For instance, for fruit classification, ANN is implemented by researchers more frequently than other algorithms. Being a deep learning algorithm, ANN automatically performs feature extraction. Three parameters that determine the performance of the algorithm are the number of layers, the number of neurons in each hidden layer, and the connections among neurons. Network size optimization is vital in order to attain a viable model in terms of learning rate and computational complexity [111]. To detect symptoms of disease and infection, SVM and RF algorithms are preferred. The reason behind the widespread use of these methods is their capability to learn from small but complex data in a simple way. Accordingly, ANN is the most popular algorithm for prediction, followed by linear regression and SVM. The ranking of algorithms is performed on the basis of usage frequency instead of precision. It should be noted that the high frequency of SVM does not indicate its best performance alone. Mainly, it is due to the appropriateness of SVM for both regression and classification problems and its reliability as a standard algorithm for performance benchmarking.

Concerning countries involved in research on oil palm, results indicate that researchers from Malaysia, Indonesia, UK, the USA, Australia, China, Thailand, and the Netherlands are actively participating. From the literature, it was also found that some prominent research and development (R&D) organizations in the world like CENIPALMA (Columbia), NIFOR (Nigeria), CIRAD (France), IOPRI (Indonesia), and MPOB (Malaysia) deal with multiple aspects of oil palm plantation development in their countries. FELDA is another local government agency in Malaysia that encourages different organizations for agricultural research [112,113]. Current development in the global and Malaysian oil palm industry is a result of significant research combined with good management, a compatible climate, and well-established infrastructure [112]. Given one full credit to each author, statistics revealed that Malaysia is ranked number 1 for publishing approximately 62% of the articles, followed by Indonesia with 18% of the articles of the reviewed material.

Investigation about data sources found that Google earth engine (GEE), QuickBird, Worldview3, as well as some other tools and repositories provided considerable satellite imagery. The vast amount of self-collected site-specific images, field samples, expert opinions, and lab test results were also well utilized. The availability of satellite images is advantageous, but it is also an indication of a massive shortage of freely available quantitative data for research in the oil palm field. Data unavailability is the main challenge for precision farming of oil palm since consistent field data collection is a time-consuming and challenging task. The findings reveal that data collection is progressing with several impediments like difficulty in quantifying yield-reducing factors and real-time data storages.

A collaborative database management system (DBMS) integrated with big data and cloud computing is entailed where oil palm observations gathered during the entire lifespan of plants should be stacked by expert stakeholders since the extensive application of ML is restrained by the availability of gold standard datasets containing fine-grained qualitative and quantitative information.

Dealing with Q7 became complex due to many validation methods. Most of the researchers validated the algorithm with multiple techniques, and performance was also compared with other existing algorithms. Some works did not clearly mention the evaluation metrics, while few did not evaluate the algorithm at all. Expert opinion and human graders were also declared as the model's performance evaluation criteria in some studies. Based on the available information, it was found that the most popular evaluation parameter was overall accuracy (OA), followed by classification accuracy (CA), root mean squared error (RMSE), and F1-score.

### *5.2. General Discussion*

Such research was inclined to risks to validity as potential threats to validity can be external, hypothetical validity, and consistency [114]. To address the aforementioned validity issues (i.e., external legitimacy and construct rationality), queries with broad initial search keywords returned a considerable number of studies. The introduced search string covered the entire scope of the proposed review. For the reliability of the review, the validity can be considered well addressed as the methodology of the review was well explained. It can be simulated as it is categorically replicable. If this review is replicated, marginally different selected publications can be returned, but the different personal judgments would be the reason behind the potential differences. However, it is very unlikely to obtain dissimilar overall findings.

### *5.3. Search-Based Discussion*

Perhaps some important publications have been intact. Different search strings could have been applied, and a wider search could have reverted with some different studies, but the satisfactory number of articles indicates that the search was broad enough to address the research questions. The most challenging part of searching for articles was the inability of search engines to differentiate between "oil palm" (tree) and "palm oil" (oil), which caused a large number of less relevant articles. Similarly, some articles containing the words "palm" (part of hand), "coconut oil," "coconut palm," and "date palm" were also searched and were discarded during article exclusion.

### *5.4. Analysis Based Discussion*

Another potential threat to validity was the method of conducting analysis. All publications did not clearly state the information required to answer our predefined research questions. In addition, strict organization of articles with more than one objective or compound methodology was not possible. For example, "regression for /and classification," "classification for disease detection," "FFB analysis for classification," or "FFB analysis for disease detection/prediction" and "UAV for trees classification" outwardly caused duplication in the grouping. However, the overall analysis remained focused on the primary objective.

### *5.5. Research Questions Based Discussion*

Some relevant studies published during our time range might have been missed because of different keywords. It was practically not possible to include every individual MLA in our search string. Deep learning and machine learning algorithms were not separated explicitly. Some of the articles were written by many authors from different countries, so the publications were calculated multiple times based on the affiliations of authors during geographical distribution analysis. Similarly, applied algorithms were calculated based on the frequency of use instead of precision. Performance comparison

of the algorithms was not considered in the results due to a variety of objectives and varied data sets, including some calculated parameters. SVM is ranked as the highest-used algorithm because of the number of studies that applied it for hybrid modelling, comparison, and performance evaluation of the proposed model. With respect to data sources, it was observed that proper distribution of the data set used was not possible due to the wide range of sources and input variables. Similarly, evaluation parameters vary in different studies. In addition, many studies did not mention evaluation parameters clearly. Most of the studies used multiple evaluation parameters, which are considered just in case of being exceptionally mentioned in work.

### 5.6. Challenges-Based Discussion

Regarding challenges, researchers discussed general challenges in the form of research motivation. Several challenges are site or data-related, while other challenges were faced during model implementation. We explained only some common challenges confined with oil palm. Another separate technical review is planned to discuss technical challenges encountered during the application of different MLAs. From the current literature, the shortage of data availability for important parameters seems the reason for the limited performance of the proposed models. When site-specific interconnected parameters are measured and included, proposed models will have better precision.

## 6. Conclusions

This study investigated existing research on the application of machine learning in oil palm actively carried out by researchers during the last decade (2011–2020). For better understanding, the reviewed articles were divided into two categories: classification and regression analysis. The study identified the key objectives in the classification category as disease detection, multipurpose classification, land cover/tree detection, fresh fruit bunch analysis, and automated canopy management/segmentation using UAV. In contrast, the regression analysis included the prediction of FFB yield, CPO prices, harvest time, and seasonal impacts on oil palm. It was evident from the results that the research was mainly focused on land cover/tree detection and oil palm plantation area observation. The other preferred research direction was fresh fruit bunch ripeness identification. Surprisingly, predictive models for fruit yield, oil production, plant growth, and disease forecasting have not been widely applied. The results clearly show that the highly applied algorithm was SVM followed by ANN, RF, regression, and CART. A deeper analysis indicated the scarcity of available data in the field. Most of the data used for oil palm research was in the form of images collected from different satellite or remote-sensing sources. Due to the limited involvement of researchers from a small number of countries, several important research dimensions were missing, such as soil classification to identify suitable land, automated pest and weed recognition, identifying the symptoms of sunlight, nutrients and water limitations in oil palm crop, optimization of fertilizers, seed assessment, and so on. Most importantly, the research trends in the literature showed that existing MLAs and techniques were not adequately coupled together to support effective decision-making systems as compared to several other domains of ML applications. Current research is too inadequate to design practically supportive tools that are capable of increasing yields and improving quality and plantation sustainability in an environmentally friendly way. Integrated ML practices such as big data, remote sensing, data analytics, image-processing, and automated information extraction are progressing to achieve knowledge-based oil palm agriculture. This study provided a vivid idea to understand the advancement of oil palm associated with machine learning and can inspire researchers to find out the relevant problems in this area. Different outcomes of this study magnify the current research on “oil palm and machine learning” from various perspectives. We believe this article will pave the way to the development of oil palm with the help of automation and intelligence.

To improve the research, future work should focus on a technical review to analyse the performance of ML models applied in the oil palm agriculture domain.



**Funding:** This work was funded by the School of Industrial Technology, Universiti Sains Malaysia, under Grant 203.PTEKIND.6777007.

**Acknowledgments:** The authors would like to thank Ministry of Higher Education, Malaysia, Long Term Research Grant Scheme for providing resources and financial support.

**Conflicts of Interest:** The authors declare no conflict of interest.

## Nomenclature

### General abbreviations

ML	Machine learning
DL	Deep learning
MLA	Machine learning algorithm
AI	Artificial intelligence
RGB	Red-green-blue
CPO	Crude palm oil
FFB	Fresh fruit bunch
RBES	Rule-based expert system
G. boninense	Ganoderma boninense
UML	Unified modeling language
SLR	Systematic literature review
UAV	Unmanned aerial vehicle
GEE	Google Earth engine
DBMS	Data base management system
GS	Genome selection
BSR	Basal stem rot
OER	Oil extraction ratio

### Abbreviations for machine learning techniques and algorithms

SVM	Support vector machine
KNN	K-nearest neighbor
ANN	Artificial neural network
CNN	Convolutional neural network
CART	Classification and regression tree
RF	Random forest
RFT	Random forest tree
DA	Discriminant analysis
LR	Logistic regression
DT	Decision tree
MA	Moving average
GDA	Gradient descent algorithm
RR	Ridge regression
GA	Genetic algorithm
MLC	Maximum likelihood classifier

### Abbreviations for machine learning models

LM	Linear model
VGG-SSD	Visual geometry group-single shot detector
HOG	Histogram of oriented gradients
LMT	Logistic model tree
ARFIMA	Autoregressive fractional integral moving average
ANFIS	Adaptive neuro-fuzzy inference system
SGBT	Stochastic gradient boosting trees (SGBT)
GANN	Genetic algorithm neural network
BLUP	Best linear unbiased prediction

### Abbreviations for model performance evaluation metrics

OA	Overall accuracy
PA	Prediction accuracy
UA	User accuracy
CA	Classification accuracy
RMSE	Root mean squared error
R <sup>2</sup>	R squared (coefficient of determination)
MAE	Mean absolute error
DA	Detection accuracy

RR	Ridge regression
DT	Decision tree
MA	Moving average
GDA	Gradient descent algorithm
LASSO	Least absolute shrinkage and selection operator
NB	Naïve Bayes
SMOTE	Synthetic minority over-sampling technique
AdaBoost	Adaptive boosting
GA	Genetic algorithm
PR	Polynomial regression
SVR	Support vector regression
PCA	Principal component analysis
JNB	Jenks natural breaks
GDA	Gaussian discriminant analysis

## References

- Singh, D.; Sharma, D.; Soni, S.L.; Inda, C.S.; Sharma, S.; Sharma, P.K.; Jhalani, A. A Comprehensive Review on 1st-Generation Biodiesel Feedstock Palm Oil: Production, Engine Performance, and Exhaust Emissions. *BioEnergy Res.* **2020**, *14*, 1–22. [\[CrossRef\]](#)
- Corley, R. How much palm oil do we need? *Environ. Sci. Policy* **2009**, *12*, 134–139. [\[CrossRef\]](#)
- Ming, K.K.; Chandramohan, D. Malaysian palm oil industry at crossroads and its future direction. *Oil Palm Ind. Econ. J.* **2002**, *2*, 10–15.
- Barcelos, E.; Rios, S.D.A.; Cunha, R.N.; Lopes, R.; Motoike, S.Y.; Babiyuchuk, E.; Skiryucz, A.; Kushnir, S. Oil palm natural diversity and the potential for yield improvement. *Front. Plant Sci.* **2015**, *6*, 190. [\[CrossRef\]](#)
- Kushairi, A.; Singh, R.; Ong-Abdullah, M. The oil palm industry in Malaysia: Thriving with transformative technologies. *J. Oil Palm Res.* **2017**, *29*, 431–439.
- Rahman, S.A.Z.; Mitra, K.C.; Islam, S.M. Soil classification using machine learning methods and crop suggestion based on soil series. In Proceedings of the 2018 21st International Conference of Computer and Information Technology (ICCIT), Dhaka, Bangladesh, 21–23 December 2018.
- Chlingaryan, A.; Sukkarieh, S.; Whelan, B. Machine learning approaches for crop yield prediction and nitrogen status estimation in precision agriculture: A review. *Comput. Electron. Agric.* **2018**, *151*, 61–69. [\[CrossRef\]](#)
- Dimitriadis, S.; Goumopoulos, C. Applying machine learning to extract new knowledge in precision agriculture applications. In Proceedings of the 2008 Panhellenic Conference on Informatics, Samos, Greece, 28–30 August 2008.
- Behmann, J.; Mahlein, A.K.; Rumpf, T.; Römer, C.; Plümer, L. A review of advanced machine learning methods for the detection of biotic stress in precision crop protection. *Precis. Agric.* **2015**, *16*, 239–260. [\[CrossRef\]](#)
- Li, W.; Fu, H.; Yu, L.; Cracknell, A. Deep learning based oil palm tree detection and counting for high-resolution remote sensing images. *Remote Sens.* **2017**, *9*, 22. [\[CrossRef\]](#)
- Santoso, H.; Tani, H.; Wang, X. Random Forest classification model of basal stem rot disease caused by *Ganoderma boninense* in oil palm plantations. *Int. J. Remote Sens.* **2017**, *38*, 4683–4699. [\[CrossRef\]](#)
- Van Dijk, M.; Morley, T.; Jongeneel, R.; van Ittersum, M.; Reidsma, P.; Ruben, R. Disentangling agronomic and economic yield gaps: An integrated framework and application. *Agric. Syst.* **2017**, *154*, 90–99. [\[CrossRef\]](#)
- Cintra, M.E.; Meira, C.A.; Monard, M.C.; Camargo, H.A.; Rodrigues, L.H. The use of fuzzy decision trees for coffee rust warning in Brazilian crops. In Proceedings of the 2011 11th International Conference on Intelligent Systems Design and Applications, Cordoba, Spain, 22–24 November 2011.
- Fan, J.; Wu, L.; Zhang, F.; Cai, H.; Zeng, W.; Wang, X.; Zou, H. Empirical and machine learning models for predicting daily global solar radiation from sunshine duration: A review and case study in China. *Renew. Sustain. Energy Rev.* **2019**, *100*, 186–212. [\[CrossRef\]](#)
- Föhse, D.; Claassen, N.; Jungk, A. Phosphorus efficiency of plants. *Plant Soil* **1988**, *110*, 101–109. [\[CrossRef\]](#)
- Kaiser, E.-A.; Kohrs, K.; Kücke, M.; Schnug, E.; Heinemeyer, O.; Munch, J.C. Nitrous oxide release from arable soil: Importance of N-fertilization, crops and temporal variation. *Soil Biol. Biochem.* **1998**, *30*, 1553–1563. [\[CrossRef\]](#)
- Cogato, A.; Meggio, F.; De Antoni Migliorati, M.; Marinello, F. Extreme weather events in agriculture: A systematic review. *Sustainability* **2019**, *11*, 2547. [\[CrossRef\]](#)
- Paul, J.; Criado, A.R. The art of writing literature review: What do we know and what do we need to know? *Int. Bus. Rev.* **2020**, *29*, 101717. [\[CrossRef\]](#)
- Chong, K.L.; Kanniah, K.D.; Pohl, C.; Tan, K.P. A review of remote sensing applications for oil palm studies. *Geo-Spat. Inf. Sci.* **2017**, *20*, 184–200. [\[CrossRef\]](#)
- Kurnia, J.C.; Jangam, S.V.; Akhtar, S.; Sasmito, A.P.; Mujumdar, A.S. Advances in biofuel production from oil palm and palm oil processing wastes: A review. *Biofuel Res. J.* **2016**, *3*, 332–346. [\[CrossRef\]](#)
- Barbedo, J.G.A. Detection of nutrition deficiencies in plants using proximal images and machine learning: A review. *Comput. Electron. Agric.* **2019**, *162*, 482–492. [\[CrossRef\]](#)
- Khosrokhani, M.; Khairunniza-Bejo, S.; Pradhan, B. Geospatial technologies for detection and monitoring of *Ganoderma* basal stem rot infection in oil palm plantations: A review on sensors and techniques. *Geocarto Int.* **2018**, *33*, 260–276. [\[CrossRef\]](#)
- Pandey, R.; Naik, S.; Marfatia, R. Image processing and machine learning for automated fruit grading system: A technical review. *Int. J. Comput. Appl.* **2013**, *81*, 29–39. [\[CrossRef\]](#)
- Rashid, M.; Bari, B.S.; Yusup, Y.; Kamaruddin, M.A.; Khan, N. A Comprehensive Review of Crop Yield Prediction Using Machine Learning Approaches with Special Emphasis on Palm Oil Yield Prediction. *IEEE Access* **2021**, *9*, 63406–63439. [\[CrossRef\]](#)
- Uning, R.; Latif, M.T.; Othman, M.; Juneng, L.; Mohd Hanif, N.; Nadzir, M.S.M.; Maulud, K.N.A.; Jaafar, W.S.W.M.; Said, N.T.S.; Ahamad, F.; et al. A review of Southeast Asian oil palm and Its CO<sub>2</sub> fluxes. *Sustainability* **2020**, *12*, 5077. [\[CrossRef\]](#)
- Von Uexkull, H.; Fairhurst, T. *Oil Palm*; International Potash Institute: Singapore, Singapore, 1992.
- Sutherland, S. What makes a weed a weed: Life history traits of native and exotic plants in the USA. *Oecologia* **2004**, *141*, 24–39. [\[CrossRef\]](#) [\[PubMed\]](#)
- Legros, S.; Mialet-Serra, I.; Caliman, J.P.; Siregar, F.A.; Clément-Vidal, A.; Fabre, D.; Dingkuhn, M. Phenology, growth and physiological adjustments of oil palm (*Elaeis guineensis*) to sink limitation induced by fruit pruning. *Ann. Bot.* **2009**, *104*, 1183–1194. [\[CrossRef\]](#)
- Ismail, A.; Mamat, M.N. The optimal age of oil palm replanting. *Oil Palm Ind. Econ. J.* **2002**, *2*, 11–18.

30. Jelsma, I.; Woittiez, L.S.; Ollivier, J.; Dharmawan, A.H. Do wealthy farmers implement better agricultural practices? An assessment of implementation of Good Agricultural Practices among different types of independent oil palm smallholders in Riau, Indonesia. *Agric. Syst.* **2019**, *170*, 63–76. [[CrossRef](#)]
31. Gérard, A.; Wollni, M.; Hölscher, D.; Irawan, B.; Sundawati, L.; Teuscher, M.; Kreft, H. Oil-palm yields in diversified plantations: Initial results from a biodiversity enrichment experiment in Sumatra, Indonesia. *Agric. Ecosyst. Environ.* **2017**, *240*, 253–260. [[CrossRef](#)]
32. Rhebergen, T.; Zingore, S.; Giller, K.E.; Frimpong, C.A.; Acheampong, K.; Ohipeni, F.T.; Panyin, E.K.; Zutah, V.; Fairhurst, T. Closing yield gaps in oil palm production systems in Ghana through Best Management Practices. *Eur. J. Agron.* **2020**, *115*, 126011. [[CrossRef](#)]
33. Alpaydin, E. *Introduction to Machine Learning*; The MIT Press: Cambridge, MA, USA, 2010.
34. Sacchelli, S.; Fabbrizzi, S.; Menghini, S. Climate change effects and adaptation strategies in the wine sector: A quantitative literature review. *Wine Econ. Policy* **2016**, *5*, 114–126. [[CrossRef](#)]
35. Butler, A.; Hall, H.; Copnell, B. A guide to writing a qualitative systematic review protocol to enhance evidence-based practice in nursing and health care. *Worldviews Evid.-Based Nurs.* **2016**, *13*, 241–249. [[CrossRef](#)] [[PubMed](#)]
36. Fakis, A.; Hilliam, R.; Stoneley, H.; Townend, M. Quantitative analysis of qualitative information from interviews: A systematic literature review. *J. Mix. Methods Res.* **2014**, *8*, 139–161. [[CrossRef](#)]
37. Shamseer, L.; Moher, D.; Clarke, M.; Ghersi, D.; Liberati, A.; Petticrew, M.; Shekelle, P.; Stewart, L.A. Preferred reporting items for systematic review and meta-analysis protocols (PRISMA-P) 2015: Elaboration and explanation. *Br. Med. J.* **2015**, *349*. [[CrossRef](#)] [[PubMed](#)]
38. Teng, S.; Khong, K.W.; Ha, N.C. Palm oil and its environmental impacts: A big data analytics study. *J. Clean. Prod.* **2020**, *274*, 122901. [[CrossRef](#)]
39. Moher, D.; Liberati, A.; Tetzlaff, J.; Altman, D.G.; Prisma Group. Preferred reporting items for systematic reviews and meta-analyses: The PRISMA statement. *PLoS Med.* **2009**, *6*, e1000097. [[CrossRef](#)]
40. Van Klompenburg, T.; Kassahun, A.; Catal, C. Crop yield prediction using machine learning: A systematic literature review. *Comput. Electron. Agric.* **2020**, *177*, 105709. [[CrossRef](#)]
41. Kwong, Q.B.; Teh, C.K.; Ong, A.L.; Chew, F.T.; Mayes, S.; Kulaveerasingam, H.; Tammi, M.; Yeoh, S.H.; Appleton, D.R.; Harikrishna, J.A. Evaluation of methods and marker Systems in Genomic Selection of oil palm (*Elaeis guineensis* Jacq.). *BMC Genet.* **2017**, *18*, 107. [[CrossRef](#)] [[PubMed](#)]
42. Amiruddin, A.D.; Muharam, F.M.; Ismail, M.H.; Ismail, M.F.; Tan, N.P.; Karam, D.S. Hyperspectral remote sensing for assessment of chlorophyll sufficiency levels in mature oil palm (*Elaeis guineensis*) based on frond numbers: Analysis of decision tree and random forest. *Comput. Electron. Agric.* **2020**, *169*, 105221. [[CrossRef](#)]
43. Amiruddin, A.D.; Muharam, F.M.; Ismail, M.H.; Tan, N.P.; Ismail, M.F. Hyperspectral spectroscopy and imbalance data approaches for classification of oil palm's macronutrients observed from frond 9 and 17. *Comput. Electron. Agric.* **2020**, *178*, 105768. [[CrossRef](#)]
44. DB, M.Y.; Rafie, A.M.; Abd Aziz, S.; Azrad, S.; Masri, M.M.M.; Shahi, A.; Marzuki, O.F. Classification of oil palm female inflorescences anthesis stages using machine learning approaches. *Inf. Process. Agric.* **2020**. [[CrossRef](#)]
45. H azir, M.H.M.; Shariff, A.R.M.; Amiruddin, M.D.; Ramli, A.R.; Saripan, M.I. Oil palm bunch ripeness classification using fluorescence technique. *J. Food Eng.* **2012**, *113*, 534–540. [[CrossRef](#)]
46. A dedayo, O.O.; Isa, M.; Che, S.A.; Abbas, Z. Comparison of Feed Forward Neural Network Training Algorithms for Intelligent Modeling of Dielectric Properties of Oil Palm Fruitlets. *Int. J. Eng. Adv. Technol. (IJEAT)* **2014**, *3*, 38–42.
47. Santoso, H.; Tani, H.; Wang, X.; Prasetyo, A.E.; Sonobe, R. Classifying the severity of basal stem rot disease in oil palm plantations using WorldView-3 imagery and machine learning algorithms. *Int. J. Remote Sens.* **2019**, *40*, 7624–7646. [[CrossRef](#)]
48. Khaled, A.Y.; Abd Aziz, S.; Bejo, S.K.; Nawi, N.M.; Seman, I.A. Spectral features selection and classification of oil palm leaves infected by Basal stem rot (BSR) disease using dielectric spectroscopy. *Comput. Electron. Agric.* **2018**, *144*, 297–309. [[CrossRef](#)]
49. Nababan, M.; Laia, Y.; Sitanggang, D.; Sihombing, O.; Indra, E.; Siregar, S.; Purba, W.; Mancur, R. The diagnose of oil palm disease using Naive Bayes Method based on Expert System Technology. *J. Phys. Conf. Ser.* **2018**, *1007*, 012015. [[CrossRef](#)]
50. Ahmadi, P.; Muharam, F.M.; Ahmad, K.; Mansor, S.; Seman, I.A. Early detection of Ganoderma basal stem rot of oil palms using artificial neural network spectral analysis. *Plant Dis.* **2017**, *101*, 1009–1016. [[CrossRef](#)] [[PubMed](#)]
51. Hunt, E.R.; Hively, W.D.; Fujikawa, S.J.; Linden, D.S.; Daughtry, C.S.; McCarty, G.W. Acquisition of NIR-green-blue digital photographs from unmanned aircraft for crop monitoring. *Remote Sens.* **2010**, *2*, 290–305. [[CrossRef](#)]
52. Fawcett, D.; Azlan, B.; Hill, T.C.; Kho, L.K.; Bennie, J.; Anderson, K. Unmanned aerial vehicle (UAV) derived structure-from-motion photogrammetry point clouds for oil palm (*Elaeis guineensis*) canopy segmentation and height estimation. *Int. J. Remote Sens.* **2019**, *40*, 7538–7560. [[CrossRef](#)]
53. Xia, M.; Li, W.; Fu, H.; Yu, L.; Dong, R.; Zheng, J. Fast and robust detection of oil palm trees using high-resolution remote sensing images. In *Automatic Target Recognition XXIX*; International Society for Optics and Photonics: Baltimore, MD, USA, 2019. [[CrossRef](#)]
54. Wang, Y.; Zhu, X.; Wu, B. Automatic detection of individual oil palm trees from UAV images using HOG features and an SVM classifier. *Int. J. Remote Sens.* **2019**, *40*, 7356–7370. [[CrossRef](#)]
55. Kalantar, B.; Idrees, M.O.; Mansor, S.; Halin, A.A. Smart counting—Oil palm tree inventory with UAV. *Coordinates* **2017**, *13*, 17–22.

56. Zorteza, M.; Nery, M.; Ruga, B.; Carvalho, L.B.; Bastos, A.C. Oil-palm tree detection in aerial images combining deep learning classifiers. In Proceedings of the IGARSS 2018-2018 IEEE International Geoscience and Remote Sensing Symposium, Valencia, Spain, 22–27 July 2018.
57. Khaki, S.; Wang, L.; Archontoulis, S.V. A cnn-rnn framework for crop yield prediction. *Front. Plant Sci.* **2020**, *10*, 1750. [[CrossRef](#)]
58. Sinambela, R.; Mandang, T.; Subrata, I.D.M.; Hermawan, W. Application of an inductive sensor system for identifying ripeness and forecasting harvest time of oil palm. *Sci. Hortic.* **2020**, *265*, 109231. [[CrossRef](#)]
59. Diana, S.R.; Purnama, S.M.; Dharma, G.; Sutrisnanto, A.; Perwitasari, I.; Farida, F. Estimation the Amount of Oil Palm Production Using Artificial Neural Network and NDVI SPOT-6 Imagery. *Int. J. Innov. Sci. Res. Technol.* **2019**, *4*, 548–554.
60. Karia, A.A.; Bujang, I.; Ahmad, I. Forecasting on crude palm oil prices using artificial intelligence approaches. *Am. J. Oper. Res.* **2013**, *3*, 29230. [[CrossRef](#)]
61. Kanchymalay, K.; Salim, N.; Sukprasert, A.; Krishnan, R.; Hashim, U.R.A. Multivariate time series forecasting of crude palm oil price using machine learning techniques. *IOP Conf. Ser.: Mater. Sci. Eng.* **2017**, *226*, 012117. [[CrossRef](#)]
62. Mustakim, M.; Buono, A.; Hermadi, I. Performance comparison between support vector regression and artificial neural network for prediction of oil palm production. *J. Ilmu Komput. Dan Inf.* **2016**, *9*, 1–8.
63. Combres, J.-C.; Pallas, B.; Rouan, L.; Mialet-Serra, I.; Caliman, J.-P.; Braconnier, S.; Soulié, J.-C.; Dingkuhn, M. Simulation of inflorescence dynamics in oil palm and estimation of environment-sensitive phenological phases: A model based analysis. *Funct. Plant Biol.* **2013**, *40*, 263–279. [[CrossRef](#)]
64. Hoffmann, M.P.; Vera, A.C.; Van Wijk, M.T.; Giller, K.E.; Oberthür, T.; Donough, C.; Whitbread, A.M. Simulating potential growth and yield of oil palm (*Elaeis guineensis*) with PALMSIM: Model description, evaluation and application. *Agric. Syst.* **2014**, *131*, 1–10. [[CrossRef](#)]
65. Zuhaimy, I.; Azme, K. Neural network in modeling Malaysian oil palm yield. *Am. J. Appl. Sci.* **2011**, *8*, 796–803.
66. Oetli, P.; Behera, S.K.; Yamagata, T. Climate based predictability of oil palm tree yield in Malaysia. *Sci. Rep.* **2018**, *8*, 2271. [[CrossRef](#)]
67. Chapman, R.; Cook, S.; Donough, C.; Lim, Y.L.; Ho, P.V.V.; Lo, K.W.; Oberthür, T. Using Bayesian networks to predict future yield functions with data from commercial oil palm plantations: A proof of concept analysis. *Comput. Electron. Agric.* **2018**, *151*, 338–348. [[CrossRef](#)]
68. Alfatni, M.S.M.; Shariff, A.R.M.; Abdullah, M.Z.; Marhaban, M.H.; Shafie, S.B.; Bamiruddin, M.D.; Saaed, O.M.B. Oil palm fresh fruit bunch ripeness classification based on rule-based expert system of ROI image processing technique results. *IOP Conf. Ser. Earth Environ. Sci.* **2014**, *20*, 012018. [[CrossRef](#)]
69. Okoro, S.U.; Schickhoff, U.; Böhner, J.; Schneider, U.A. A novel approach in monitoring land-cover change in the tropics: Oil palm cultivation in the Niger Delta, Nigeria. *DIE ERDE–J. Geogr. Soc. Berl.* **2016**, *147*, 40–52.
70. Zheng, J.; Li, W.; Xia, M.; Dong, R.; Fu, H.; Yuan, S. Large-scale oil palm tree detection from high-resolution remote sensing images using faster-rcnn. In Proceedings of the IGARSS 2019, 2019 IEEE International Geoscience and Remote Sensing Symposium, Yokohama, Japan, 28 July–2 August 2019.
71. Rizeei, H.M.; Shafri, H.Z.; Mohamoud, M.A.; Pradhan, B.; Kalantar, B. Oil palm counting and age estimation from WorldView-3 imagery and LiDAR data using an integrated OBIA height model and regression analysis. *J. Sens.* **2018**, *2018*, 2536327. [[CrossRef](#)]
72. Srestasathiern, P.; Rakwatin, P. Oil palm tree detection with high resolution multi-spectral satellite imagery. *Remote Sens.* **2014**, *6*, 9749–9774. [[CrossRef](#)]
73. Li, W.; Dong, R.; Fu, H.; Yu, L. Large-scale oil palm tree detection from high-resolution satellite images using two-stage convolutional neural networks. *Remote Sens.* **2019**, *11*, 11. [[CrossRef](#)]
74. Descals, A.; Wich, S.; Meijaard, E.; Gaveau, D.L.; Peedell, S.; Szantoi, Z. High-resolution global map of smallholder and industrial closed-canopy oil palm plantations. *Earth Syst. Sci. Data Discuss.* **2020**, 1–22. [[CrossRef](#)]
75. Shaharum, N.S.N.; Shafri, H.Z.M.; Ghani, W.A.W.A.K.; Samsatli, S.; Prince, H.M.; Yusuf, B.; Hamud, A.M. Mapping the spatial distribution and changes of oil palm land cover using an open source cloud-based mapping platform. *Int. J. Remote Sens.* **2019**, *40*, 7459–7476. [[CrossRef](#)]
76. Mathew, L.S.; Spannagl, M.; Al-Malki, A.; George, B.; Torres, M.F.; Al-Dous, E.K.; Al-Azwani, E.K.; Hussein, E.; Mathew, S.; Mayer, K.F. A first genetic map of date palm (*Phoenix dactylifera*) reveals long-range genome structure conservation in the palms. *BMC Genom.* **2014**, *15*, 1–10.
77. Mubin, N.A.; Nadarajoo, E.; Shafri, H.Z.M.; Hamedianfar, A. Young and mature oil palm tree detection and counting using convolutional neural network deep learning method. *Int. J. Remote Sens.* **2019**, *40*, 7500–7515. [[CrossRef](#)]
78. Juman, M.A.; Wong, Y.W.; Rajkumar, R.K.; Goh, L.J. A novel tree trunk detection method for oil-palm plantation navigation. *Comput. Electron. Agric.* **2016**, *128*, 172–180. [[CrossRef](#)]
79. Morel, A.C.; Saatchi, S.S.; Malhi, Y.; Berry, N.J.; Banin, L.; Burslem, D.; Nilus, R.; Ong, R.C. Estimating aboveground biomass in forest and oil palm plantation in Sabah, Malaysian Borneo using ALOS PALSAR data. *For. Ecol. Manag.* **2011**, *262*, 1786–1798. [[CrossRef](#)]
80. Gutiérrez-Vélez, V.H.; DeFries, R. Annual multi-resolution detection of land cover conversion to oil palm in the Peruvian Amazon. *Remote Sens. Environ.* **2013**, *129*, 154–167. [[CrossRef](#)]



81. Abdani, S.R.; Zulkifley, M.A. Densenet with spatial pyramid pooling for industrial oil palm plantation detection. In Proceedings of the 2019 International Conference on Mechatronics, Robotics and Systems Engineering (MoRSE), Bali, Indonesia, 4–6 December 2019.
82. Shaharum, N.S.N.; Shafri, H.Z.M.; Ghani, W.A.W.A.K.; Samsatli, S.; Yusuf, B.; Al-Habshi, M.M.A.; Prince, H.M. Image classification for mapping oil palm distribution via support vector machine using Scikit-learn module. *Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci.* **2018**, *42*, 133–137. [[CrossRef](#)]
83. Shaharum, N.S.N.; Shafri, H.Z.M.; Ghani, W.A.W.A.K.; Samsatli, S.; Al-Habshi, M.M.A.; Yusuf, B. Oil palm mapping over Peninsular Malaysia using Google Earth Engine and machine learning algorithms. *Remote Sens. Appl. Soc. Environ.* **2020**, *17*, 100287. [[CrossRef](#)]
84. Li, W.; Fu, D.; Su, F.; Xiao, Y. Spatial–Temporal Evolution and Analysis of the Driving Force of Oil Palm Patterns in Malaysia from 2000 to 2018. *ISPRS Int. J. Geo-Inf.* **2020**, *9*, 280. [[CrossRef](#)]
85. Nooni, I.K.; Duker, A.A.; Van Duren, I.; Addae-Wireko, L.; Osei Jnr, E.M. Support vector machine to map oil palm in a heterogeneous environment. *Int. J. Remote Sens.* **2014**, *35*, 4778–4794. [[CrossRef](#)]
86. Culman, M.; de Farias, C.M.; Bayona, C.; Cruz, J.D.C. Using agrometeorological data to assist irrigation management in oil palm crops: A decision support method and results from crop model simulation. *Agric. Water Manag.* **2019**, *213*, 1047–1062. [[CrossRef](#)]
87. Rueda, C.; Miserque, J.; Laverde, R. Validation of an oil-palm detection system based on a logistic regression model. In Proceedings of the 2016 IEEE ANDESCON, Arequipa, Peru, 19–21 October 2016.
88. Fadilah, N.; Saleh, J.M.; Ibrahim, H.; Halim, Z.A. Oil palm fresh fruit bunch ripeness classification using artificial neural network. In Proceedings of the 2012 4th International Conference on Intelligent and Advanced Systems (ICIAS2012), Kuala Lumpur, Malaysia, 12–14 June 2012.
89. May, Z.; Amaran, M. Automated oil palm fruit grading system using artificial intelligence. *Int. J. Eng. Sci.* **2011**, *11*, 30–35.
90. Bensaeed, O.M.; Shariff, A.M.; Mahmud, A.B.; Shafri, H.; Alfatni, M. Oil palm fruit grading using a hyperspectral device and machine learning algorithm. *IOP Conf. Ser. Earth Environ. Sci.* **2014**, *20*, 012017. [[CrossRef](#)]
91. Makky, M.; Soni, P. Development of an automatic grading machine for oil palm fresh fruits bunches (FFBs) based on machine vision. *Comput. Electron. Agric.* **2013**, *93*, 129–139. [[CrossRef](#)]
92. Albakri, Z.M.; Kassim, M.S.M.; Abdullah, A.F.; Harith, H.H. Analysis of Oil Palm Leaf Phyllotaxis towards Development of Models to Determine the Fresh Fruit Bunch (FFB) Maturity Stages, Yield and Site-Specific Harvesting. *Pertanika J. Sci. Technol.* **2019**, *27*, 659–672.
93. Taparugssanagorn, A.; Siwamogsatham, S.; Pomalaza-Ráez, C. A non-destructive oil palm ripeness recognition system using relative entropy. *Comput. Electron. Agric.* **2015**, *118*, 340–349. [[CrossRef](#)]
94. Septiarini, A.; Hamdani, H.; Hatta, H.R.; Anwar, K. Automatic image segmentation of oil palm fruits by applying the contour-based approach. *Sci. Hortic.* **2020**, *261*, 108939. [[CrossRef](#)]
95. Shabdin, M.K.; Shariff, A.R.M.; Johari, M.N.A.; Saat, N.K.; Abbas, Z. A study on the oil palm fresh fruit bunch (FFB) ripeness detection by using Hue, Saturation and Intensity (HSI) approach. *IOP Conf. Ser. Earth Environ. Sci.* **2016**, *37*, 12039. [[CrossRef](#)]
96. Sabri, N.; Ibrahim, Z.; Syahlan, S.; Jamil, N.; Mangshor, N.N.A. Palm oil fresh fruit bunch ripeness grading identification using color features. *J. Fundam. Appl. Sci.* **2017**, *9*, 563–579. [[CrossRef](#)]
97. Fadilah, N.; Mohamad-Saleh, J. Color feature extraction of oil palm fresh fruit bunch image for ripeness classification. In Proceedings of the 13th International Conference on Applied Computer Science 2014, Kuala Lumpur, Malaysia, 23–25 April 2014; pp. 51–55.
98. Hazir, M.H.M.; Shariff, A.R.M.; Amiruddin, M.D. Determination of oil palm fresh fruit bunch ripeness—Based on flavonoids and anthocyanin content. *Ind. Crops Prod.* **2012**, *36*, 466–475. [[CrossRef](#)]
99. Fadilah, N.; Mohamad-Saleh, J.; Abdul Halim, Z.; Ibrahim, H.; Ali, S.S.S. Intelligent color vision system for ripeness classification of oil palm fresh fruit bunch. *Sensors* **2012**, *12*, 14179–14195. [[CrossRef](#)]
100. Harun, N.H.; Misron, N.; Sidek, R.M.; Aris, I.; Ahmad, D.; Wakiwaka, H.; Tashiro, K. Investigations on a novel inductive concept frequency technique for the grading of oil palm fresh fruit bunches. *Sensors* **2013**, *13*, 2254–2266. [[CrossRef](#)]
101. Adedayo, O.O.; Isa, M.M.; Soh, A.C.; Abbas, Z. Multi-adaptive neuro-fuzzy inference system for dielectric properties of oil palm fruitlets. *Int. J. Appl. Sci. Eng.* **2014**, *12*, 1–8.
102. Fahmi, F.; Suherman, S. Oil palm fresh fruit bunch ripeness classification using back propagation and learning vector quantization. *IOP Conf. Ser. Mater. Sci. Eng.* **2018**, *434*, 12066. [[CrossRef](#)]
103. Kassim, M.S.M.; Ismail, W.I.W.; Ramli, A.R.; Bejo, S.K. Oil palm fresh fruit bunches (FFB) growth determination system to support harvesting operation. *J. Food Agric. Environ.* **2012**, *10*, 620–625.
104. Patkar, G.; Anjaneyulu, G.S.G.N.; Mouli, P.C. Palm fruit harvester algorithm for *elaeis guineensis* oil palm fruit grading using UML. In Proceedings of the 2015 IEEE International Conference on Computational Intelligence and Computing Research (ICIC), Madurai, India, 10–12 December 2015.
105. Ibrahim, Z.; Sabri, N.; Isa, D. Palm oil fresh fruit bunch ripeness grading recognition using convolutional neural network. *J. Telecommun. Electron. Comput. Eng. (JTEC)* **2018**, *10*, 109–113.
106. Alfatni, M.S.M.; Shariff, A.R.M.; Bejo, S.K.; Saeed, O.M.B.; Mustapha, A. Real-time oil palm FFB ripeness grading system based on ANN, KNN and SVM classifiers. *IOP Conf. Ser. Earth Environ. Sci.* **2018**, *169*, 012067. [[CrossRef](#)]

107. Silalahi, D.D.; Reaño, C.E.; Lansigan, F.P.; Panopio, R.G.; Bantayan, N.C. Using genetic algorithm neural network on near infrared spectral data for ripeness grading of oil palm (*Elaeis guineensis* Jacq. ) *fresh fruit*. *Inf. Process. Agric.* **2016**, *3*, 252–261. [[CrossRef](#)]
108. Ahirwal, M.K.; Kumar, A.; Singh, G.K. Prediction and Classification. In *Computational Intelligence and Biomedical Signal Processing*; Springer: Cham, Germany, 2021; pp. 83–111. [[CrossRef](#)]
109. Braga-Neto, U. *Fundamentals of Pattern Recognition and Machine Learning*; Springer: Cham, Germany, 2020. [[CrossRef](#)]
110. Saiz-Rubio, V.; Rovira-Más, F. From smart farming towards agriculture 5.0: A review on crop data management. *Agronomy* **2020**, *10*, 207. [[CrossRef](#)]
111. Saritas, M.M.; Yasar, A. Performance analysis of ANN and Naive Bayes classification algorithm for data classification. *Int. J. Intell. Syst. Appl. Eng.* **2019**, *7*, 88–91. [[CrossRef](#)]
112. Soh, A.C.; Mayes, S.; Roberts, J.; Rajanaidu, N.; Din, A.M.; Marhalil, M.; Norziha, A.; Meilina, O.-A.; Fadila, A.M.; Azwani, A.B.N.; et al. Genetic Resources. In *Oil Palm Breeding: Genetics and Genomics*; CRC Press: Boca Raton, FL, USA, 2017; pp. 19–56.
113. Rethinam, P.; Murugesan, P. Global perspective of germplasm and breeding for seed production in oil palm. *Int. J. Oil Palm* **2018**, *10*, 17–34.
114. Feldt, R.; Magazinius, A. Validity threats in empirical software engineering research—An initial survey. In Proceedings of the 22nd International Conference on Software Engineering & Knowledge Engineering (SEKE'2010), Redwood City, San Francisco Bay, CA, USA, 1–3 July 2010; pp. 374–379.