Artificial Intelligence (AI) and Machine Learning (ML) tools for image driven 2D materials (e.g., Graphene) discovery

Hafiz Munsub Ali, Venkata Gadhamshetty, Bharat Jasthi, Etienne Z. Gnimpieba

Abstract:

Objective: The increased interest in 2D material development lies in its ability to address diverse problems in various domains, from infrastructure to biomedicine applications. Recent advances in data science and data driven discovery allow scientist to leverage artificial intelligence (AI) and machine learning (ML) techniques to tackle complex scientific challenges involved 2D material. Among various data modalities used in 2D material discovery, microscopic and spectroscopic data present a unique capability and involved in almost every developmental process from material design to characterization and maintenance. This paper presents a systematic review of current advances, issues, research gaps, and future trends in 2D material based microscopic and spectroscopic data using AI and ML techniques. ***Method***: We use PRISMA, an objective design method for literature reviews, for this review, whereas PRISMA relies on a rigorous process to improve the relevance of the selected papers and the reproducibility of the review process. Our design allowed us to query three publication databases (WoS, PubMed, Dimensions) and focus on two key research questions. The Quality of data in relevant papers was assessed using the FAIR (Findable, Accessible, Interoperable, Reusable) compliance criteria to evaluate and encourage reproducibility and shareability of research outcomes. ***Results:*** PRISMA identify 31 key articles relevant to this topic. Since 2017, an upward trend in appreciation of AI and ML techniques for the tasks such as prediction, classification, clustering, deep learning, and AI, to solve the issues involve 2D materials using microscopic images and spectroscopic data. From the relevant set of articles, the tasks prediction, classification, deep learning, clustering, and AI appeared 38.98%, 28.81%, 13.56%, 10.17%, and 8.47% respectively. Here, AI task refers to the combination of two or more ML methods or the combination of one ML and one other method to solve the data driven issues. Current research trend (i.e., 8.47%) shows that AI is still an underexplored computing technique. ***Conclusion:*** The most studied questions are 2D new material synthesis and engineering with reasonable appreciation for functional discovery/property, defect characterization and grain boundary, but work on corrosion application/detection is not done yet as per our analysis. Diversity and range of microscopic and spectroscopic data implies multi-level difficulty due to: (a) types of apparatus to generate data, (b) availability of computational resources and domain knowledge, (c) AI and ML techniques to solve specific problems. A lack in homogenous validation criteria for different representation and identification methods is observed which is an obstacle in standardized performance comparisons for different AI/ML algorithms. ***Significance***: This review is envisioned for broader research community who may be interested to focus their research on following three keywords: "*2D materials*", "*microscopic mage and spectroscopic data*", "*AI and ML*".

Keywords: 2D materials, microscopic images, spectroscopic data, artificial intelligence (AI), machine learning (ML),

Table of Contents

1. Introduction

Since invention of graphene in 2004, developments and applications of two-dimensional (2D) materials has been revolutionizing material sciences and its allied areas of interdisciplinary research. Its far-reaching impact can be witnessed in the areas like biomedical applications [1], material coatings [2] [3], electronics [4], etc. For example, the development of 2D monoelemental nanomaterials (Xenes) has generated enormous interest in biomedical applications due to its better optical and electronic features. Xenes as biological theranostic agents may address various challenges faced by the modern healthcare [1]. Similarly impacts and interactions of 2D coatings on biological matters are discussed in [2] [3] while superior features of 2D materials open the doors of innovation in electronics [4]. Vide spread applications of 2D materials is a source of motivation for researcher to discover new 2D materials, and current research on discovering knowledge through exploring structural, thermodynamic, elastic, electronic, magnetic, and optical properties of these materials lead to further innovative applications. A Computational 2D Materials Database (C2DB) hosts data for around 4000 2D materials [5] while [1] [2] [3] [4] [6] shows the focus of current research and range of 2D material applications.

1.1 Overview of 2D materials challenges and tools

Here, we briefly discuss challenges faced by the applications of 2D materials and potential AI and ML tools to solve these challenges (Figure 1). Like several other disciplines, recently artificial intelligence (AI) and machine learning (ML) techniques have been a focus of interdisciplinary research community working towards material informatics, specifically to reinforce and expedite their wet lab discoveries. ML is a data-driven (text, images, etc.) computing paradigm while AI is an umbrella term used for the combination of ML methods, including applied and natural sciences knowledge like evolutionary computation, mathematics, statistics, etc. The goal of the AI is to develop a machine (or software agent) that learn from past experiences. A common example in this regard is facial, fingerprint or object recognition systems. Images and thumb impression data is being used as a learning sources in these systems [7]. In this context, AI and ML techniques offer a rich computing intellect to solve material related sensitive issues (Figure 1a) [8].

1.1.1 2D material challenges

Research community is facing multi-faceted challenges, such as dry/wet lab, while dealing with 2D materials since its inception in 2004. The discovery of new 2D materials, or new properties of existing 2D materials can be an exciting intellectual challenge towards improving our environment and economy Complexity of these challenges grew exponentially with the growth in their databases [9] and applications. Demystifying crystallographic structure, synthesis, corrosion, and defects engineering knowledge about 2D materials help customizing applications with desired features that are not only cost effective but also environment friendly. The challenges listed in the Figure 1a are nor exhaustive.

1.1.2 2D materials microscopic/spectroscopic data

Algorithms need input data (or information) to solve 2D material problems as shown in the Figure 1 (b) and (c). Current literature shows that different types of data can be used to solve these problems. Here, we limit the scope to the datasets based on optical/spectroscopic data of 2D materials. Raw data need significant scientific/computational treatment before making it useful as input to computational tools (See

Figure 1b). After cleaning, next step is to annotate small number of images as compared to total number of available images. Annotation is a manual (or semi-manual using software) labeling of images or objects within images and that can be saved as json/csv for input to the algorithms. Preprocessing step can improve the image by suppressing noisy information and improve images to enhance relevant features to produce better output [10]. Feature engineering is the process of extracting, selecting, and constructing features with respect to domain knowledge of the problem which is used as a notion of learning during the AI and ML algorithm operations [11]. All preprocessed data/images are input to algorithm of choice (Figure 1b) that define the problem statement and corresponding tasks.

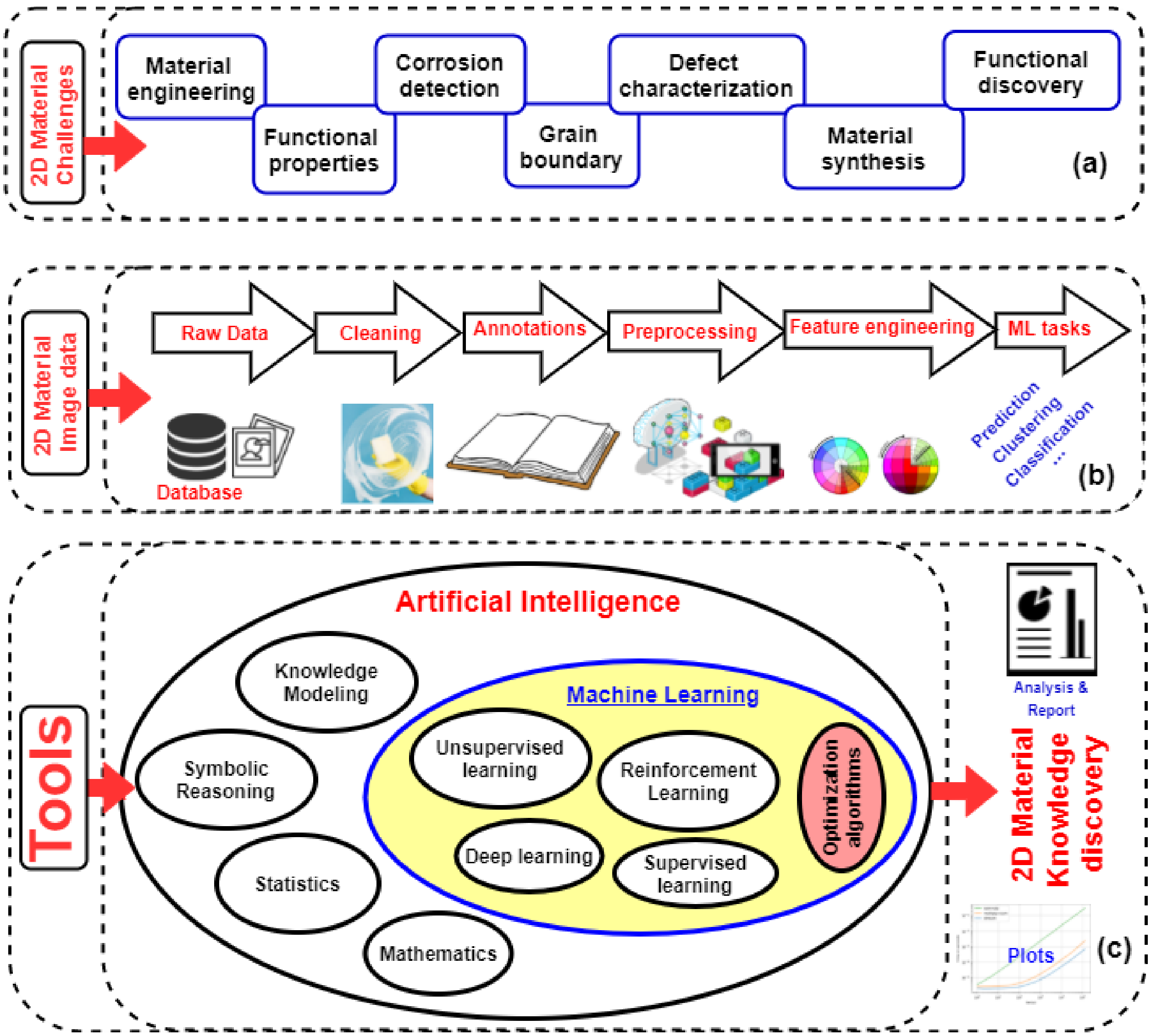


Figure 1 Overview of the systems 2D materials (a) challenges (b) with image-driven data (c) using AI and ML tools emphasizing the information-flow and interwinding nature of the subject matter in relationship to techniques used in the review papers.

1.1.3 AI Tools for study of 2D materials

Recent trends for discovery, design, and characterization of new materials are guided by data-driven computational algorithms including AI, ML, data mining (DM), etc. The term AI was coined in 1950s and is currently used for the combination of two or more ML, DM, statistical and mathematical techniques to solve material problems [11].

Artificial intelligence (AI)

Goal of AI is to develop an automatic soft machine (or software agent) for various tasks involved in 2D materials. In this review, AI refers to the combination of ML algorithm with support of other allied techniques (or sciences) such as statistics and mathematics to solve 2D material issues. For example, number of layers for 2D material are determined [12] using combination of two methods: Fresnel law and machine learning. Another example in [13], where genetic algorithms (GA) along with ML have been used to find structural defects of 2D transition-metal dichalcogenides (TMDs).

Machine learning (ML)

ML is a data-driven computing paradigm which can predict, classify, or cluster required objects accurately without explicit programing and search guidance (Figure 2). In ML, search guidance to explore huge search space may be inferred from available data and heuristics using efficient computation. Broadly, ML algorithms can be categorized into two parts: supervised and unsupervised.

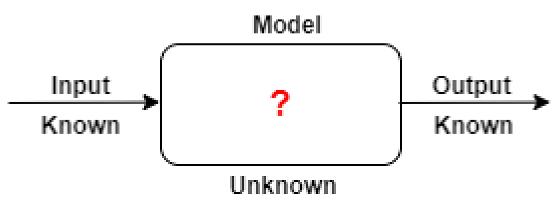


Figure 2 Machine learning computing paradigm.

Supervised ML

Supervised ML involves a series of functions that maps a set of input to the set of output based on series of identical available examples of input-output pairs. Supervised ML further divided into regression and classification. In regression, continuous target value is computed using relationship between the dependent and independent variables such as linear regression, where linear regression is used to find a best fit curve. On the other hand, target is a discrete value in classification with a countably finite number of outcomes. Some of the famous classification algorithms include logistic regression, support vector machine (SVM), k-nearest neighbors (KNN), etc.

Unsupervised ML

Unsupervised ML is subdivided into clustering and dimensionality reduction. Clustering involves grouping of data points into homogeneous groups and make sure data points in different groups should not be similar. Some of the famous algorithms are k-means, hierarchical, mean shift, etc. with a common goal to form homogenous clusters. Dimensionality reduction, feature selection, and feature extraction are also important tasks for efficient use of computational resource. Some well-known algorithms to perform these tasks include Principal Component Analysis (PCA), Independent Component Analysis (ICA), etc.

1.2 Microscopic/spectroscopic data driven techniques for 2D materials

Advances in computational intelligence and availability of relatively low-cost computational resources popularize the microscopic/spectroscopic data processing in material sciences. Recent literature shows this computational leverage to solve complex material issues with more accuracy. Inherent characteristics of 2D materials not only increase the application sensitivity but also increase the computational complexity of the problem. However, the emergence of artificial intelligence (AI) and machine learning (ML) algorithms leverage to perform several computational tasks (such as prediction, classification, clustering, etc.) using microscopic/spectroscopic data of 2D materials. This could not only reduce human labor but also improve analysis, efficiency, and accuracy [12].

Accurately finding thickness of a 2D material is a computationally tedious task. There are some commonly used data modalities to accurately gauge the thickness of 2D materials: atomic force microscopy (AFM), Raman spectroscopy, scanning electron microscopy (SEM) and high-resolution transmission electron microscopy (HRTEM). Atomic force microscopy (AFM) is a time-consuming method with limited scan range; therefore, it is not suitable to determine thickness on a large area. Raman modes are thickness dependent, and these modes have limited capability of distinguishing thickness of few-layers for 2D materials. In contrast, optical microscopy is an efficient and widely applied method to determine thickness of 2D materials over a large area. The goal of this review is to explore the performance and research gaps while using AI and ML techniques in analyzing 2D materials using microscopic/spectroscopic data [12], [14].

1.3 2D material applications

1.3.1 Biomedical applications

Emerging 2D materials have unique properties from 100 nm to several microns with one or a few atomic thicknesses, which leads to advance numerous biomedical applications. Recently developed 2D monoelemental nanomaterials (Xenes) possesses optical and electronic properties, which can act as promising biological theranostic agents to address various healthcare issues. These issues include diagnosis agents for computed tomography (CT), photoacoustic imaging (PAI), fluorescence imaging (FI), etc. and disease phototherapy such as photothermal therapy (PTT) and photodynamic therapy (PDT) against tumor, bacteria, and virus. Moreover, Xenes can build biosensors with the attachment of various biological markers like DNA, etc. to be used in sensitive healthcare applications [1]. Graphene is type of 2D material with a honeycomb structure which has numerous biomedical applications. In addition to be used in tissue engineering, gene therapy, cell imaging, and bioelectronics, graphene nanomaterials also provide some biomedical devices such as deep brain stimulators and blood glucose sensors [15][16]. Similarly, properties like intermolecular interactions of graphene variants enable the fabrication of number of functional hydrogels for biomedical applications [17]. Further, graphene-based bioactive coatings on metallic biomedical devices used in human body may improve various implanting techniques through creating certain level of corrosion resistance. All these examples not only demonstrate the range of current biomedical applications using 2D materials but also show its potential in future applications.

1.3.2 2D coatings and composites

Abiotic and biotic corrosion issues not only impact negatively but also incur huge monetary costs in various domains such as military applications, the marine sector, the oil industry, utilities, pipelines, transformers, printed circuit boards, and mechanical equipment, etc. 2D hexagonal boron nitride (hBN) thin coatings shows resilience under the influence of harsh chemicals, microbes, and heat due to its weak coupling to the environment and chemical inertness properties [18]. Another maleic-anhydride-functionalized graphene nanofillers is designed to enhance corrosion resistance of epoxy coating (MAGE) on mild steel surfaces. After applications of this coating, corrosion resistance of steel increased by 9–10 orders of magnitude as compared to bare metal in both abiotic and aggressive microbial environments [19].

Hybrid graphene particles are used as composites for several critical applications like metallic biomedical devices [20], etc. Nanotechnology of 2D materials have demonstrated a novel and more complex composite materials with desired engineered properties and optimized performance [21]. Here are few examples of 2D material composites: (1) graphene-derived particles can be integrated into a polymer matrix to augment tissue regeneration. This will improve physical, chemical, and biological properties of the polymer composite. Improved mechanical and electrical properties of graphene particles can be utilized to provide mechanical support in load bearing tissues and electrical stimuli to cells, respectively [22] [23]. (2) the smart integration of the two types of functional materials (Metal–organic frameworks (MOFs) and 2D materials) can improve performance in molecular absorption, separation, and storage, with promise in selective catalysis and biosensing [24]. So, 2D coatings and 2D composites are the extension of 2D material applications.

1.3.3 Electronics

Crystallinity, impurity, and defects defines the quality of 2D materials, which may be considered as a cornerstone for fabrication of high performance 2D devices. 2D semiconducting materials have been used as building blocks in various electronic and optoelectronic devices, including transistors, flash memories, photodetectors, and light emitters. Other 2D semiconducting materials, such as black phosphorous (or phosphorene) and silicene, have also emerged as alternatives, showing high mobility [25]. Similarly, graphene-based transistors and semiconductors has created the current generation of sophisticated electronics products including smart electronic devices, bendable smart mobiles, thin and efficient circuitry, and omnipresence of touch screens [6], [26].

1.3.4 Energy

Current environmental issues are the direct results of the way energy produced and consumed. These issues include water pollution, air pollution, climate change, thermal pollution, etc. Integration of graphene into commercial electrodes is an essential step in the production of devices. Currently, synthesis and assembly of graphene into macrostructures that ranges from 0D quantum dots to potentially 4D self-folding materials. This allows the design of batteries and supercapacitors with advance features beyond the current technology. Graphene increases the lifespan of the batteries with features like charging quicker and holding more power for longer. In addition, graphene-based batteries are lighter in weight, and smaller in size. Graphene supercapacitors saved certain amount of energy consumption, and has potential to mitigate the adverse environmental impacts such as carbon emission [2], [27]. Graphene-based electrodes are also widely used in fuel cell and microbial fuel cell applications [28].

1.3.5 Graphene membranes

Shortage of purified drinking water is becoming an alarming issue with current population growth and rate of environmental pollution. Significant number of resources are being used on individual and government levels to solve this issue and its related consequences in health, and other areas. Graphene membranes technology has potential to solve water purification issues. Graphene nanomaterials has unique physicochemical properties, and it offers mechanical durability, atomic thickness, nanosized pores and reactivity toward polar and non-polar water pollutants. Their nanosized pores can effectively separate organic solvent from water and remove water from a gas mixture to an exceptional level. In addition, graphene oxide membranes can form a perfect barrier when dealing with liquids and gases. Graphene membranes can remove industrial carbon dioxide release into the atmosphere, gas separation, and water desalination [6], [29].

1.3.6 Sensors

Properties of the graphene material are useful to sensing applications, which leads the development of various types of graphene-based electronic sensors such as biological, mechanical, gas and chemical sensors. every atom in Graphene can be exposed to its surrounding environment allowing it to sense changes, for example the goal of the chemical sensors is to detect just one molecule of a potentially dangerous substance. Graphene-based micrometer-size sensors can detect individual events on a molecular level. These features enable their uses in numerous real applications, such as crop protection and monitoring, healthcare, etc. [6], [30].

1.4 Motivation and goals of the present review

The goal of this review is to systematically (using PRISMA) explore ongoing research in 2D materials that involved AI and ML techniques using microscopic/spectroscopic data as input to discover knowledge. The abstract Figure 1 illustrate the focus of this review in three parts: Figure 1a a shows the wide range of issues faced by research involved 2D materials, data modalities and potential tasks shows in Figure 1b and current/potential tools used to solve the current problems shows in Figure 1c. Exiting reviews also lacked in combination of keywords: "*2D materials*", "microscopic/spectroscopic data", "*machine learning (ML) and artificial intelligence (AI)*". A review in [11] explored methodologies used for new 2D materials discovery, design, properties and prediction using ML techniques without focusing on specific data modality. In [31], a predictive and diagnostics sensing devices are explored using AI and ML that can capture nanofibers. Nonlinear optics (NLO) spectroscopy data is used in characterizing the physical and chemical properties of 2D materials [32]. A review in [33] focused on defect topology and density in 2D materials. To the best our knowledge, no comprehensive dedicated review is available to focus on following three keywords: 2D materials, microscopic/spectroscopic data, machine learning (ML) and artificial intelligence (AI). This review focusses on AI and ML techniques that are using microscopic image and spectroscopic data to discover knowledge for 2D materials. The goal of this review is to explore shortcomings, potential loopholes, and research gaps in the current research and intend to devise future trends for upcoming research.

In the rest of the paper, the PRISMA workflow is discussed in Section 2, and the results are presented in Section 3. The discussion on these results is presented in Section 4, and finally, the paper is concluded in Section 5.

2. Methods (PRISMA)

2.1 PRISMA overview

This review is designed by using Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) workflow as shown in Figure 3. PRISMA is an objective design method for literature reviews of any field including bioscience, biomedicine, biomaterial, or material sciences. PRISMA rigorously checks and improve process and relevancy to select research articles across the databases and ensure reproducibility of the review process [34]. This review uses three publication databases, including *Web* *of Science*, *PubMed*, *Dimension*, to query research questions designed for this review (Table 1). This process yielded 1259 articles related to associated queries, as shown in Table 2. An additional article was added from other journals which was not available in the collection from the databases. After removing duplicates in the collected set of articles, we obtained 258 unique articles. Following manual curation, we retained 51 qualified articles. For example, an article that was retrieved from the searched databases because it mentions ''2D", "materials", "AI or ML" but either it was not dealing with 2D materials or it was not using microscopic/spectroscopic data as input, was excluded from the final list, (Table 3). A deeper assessment followed by discussion with domain experts/coauthors allow us to select 31 articles for this review (Table 3).

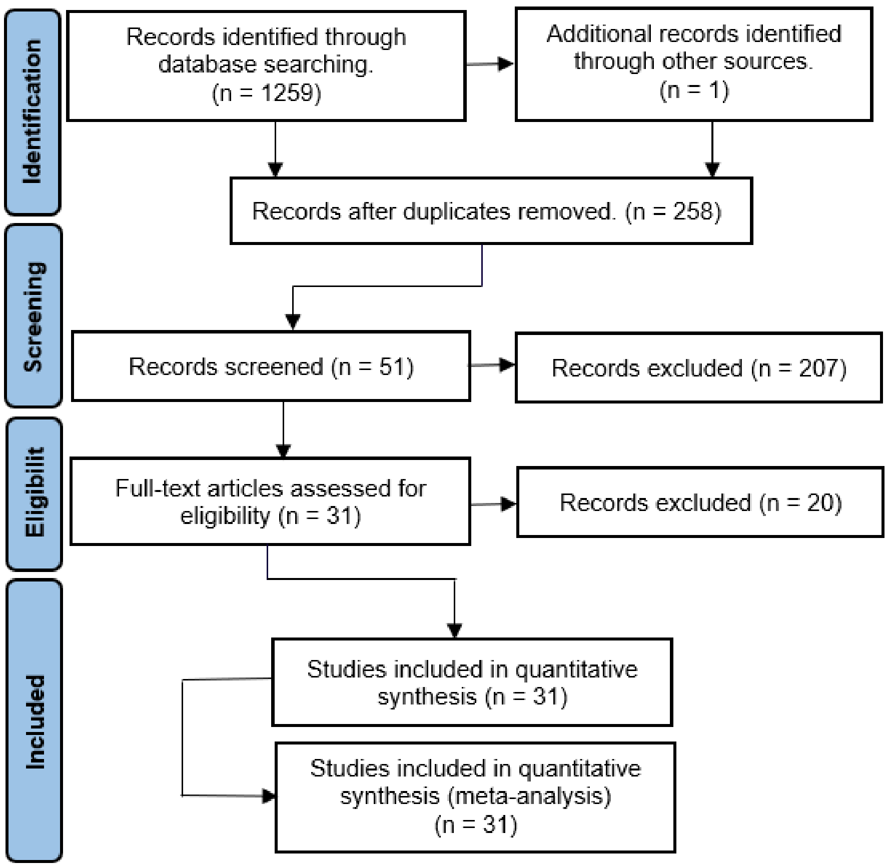


Figure 3 Search workflow (PRISMA) used in article searching, retrieving, processing and inclusion/exclusion decision making.

2.2 Identification of research questions

In this review, we were motivated by the organization of a similar article on the investigation of vector-host-pathogen relationships using data mining and machine learning [34]. Here, we undertake a rigorous effort to identify studies that used artificial intelligence (AI) and machine learning (ML) techniques to solve issues related to 2D materials and develop 2D material applications. Figure 1 shows an abstract level of knowledge discovery process which identify 2D material issues, existing data or knowledge, and potential computational methodologies to solve these issues. We started from a general problem statement and then include domain keywords, familiar domain terminologies, similar or equivalent domain keywords and terms, and methodologies. From the main research question, which we denoted as Q0 – ''To what extent Image Driven ML tools have been applied to study 2D material development and analysis?", seven subsequent research questions were formulated to integrate the systemic aspects of the problems (Table 1 and Figure 1). These research questions were mapped to 5 AI and ML tasks to better understand the systemic depth of ML tools involved in this ecosystem during the knowledge discovery processes (Table 1 and Figure 1). Focusing on the 2 most pertinent, well-defined questions (Q0, Q1), we generated 10 tasks dependent questions, bringing our research question list to a total of 18 questions. These 18 research questions were used to build our queries for database and journal searches. Based on our research questions, we crafted 8 main queries and 10 nested queries to the first 2 questions and artificial intelligence (AI) and machine learning (ML) tasks. For example, machine learning (ML) and classification form a query of ''Image Driven 2D Material AND (Classification) AND (Machine Learning)." A full table of queries is shown below (Table 2).

2.3 Search process design and selection

To perform the actual search process, queries in Table 2 were entered into the databases of Web of Science, PubMed, Dimension (June/1/ 2021) one at a time to obtain journal articles, and other publications relevant to our research questions as in (Table 1). With caution not to leave out any important articles, no date or time constraints were applied. After a successful search, the resulting articles were manually inspected using blinded manual curation to identity their relevance to 2D material applications and use of artificial intelligence (AI) and machine learning (ML) techniques. For example, Q0, was manually inspected to exclude articles that did not contain application of machine learning (ML) or artificial intelligence (AI) methods on 2D material applications with microscopic/spectroscopic data as input. Articles or studies that are not related to any of three "*AI or ML*", "*2D materials*", and "*microscopic/spectroscopic data*" were excluded. Also, manual curation was done to identify and merge overlaps resulting from searches of different databases. For instance, if articles were already indexed by PubMed, they were merged with articles from WoS (or Dimension) to avoid the duplication (Table 2.).

Table 1 Research questions mapped to AI and ML tasks. The questions are denoted as Q0 to Q7. Q0 is the main research question involving the overall relationship, Q1 to Q7 are applied to the subtle questions in the relationship with more attention to Q1 in this review. Tasks are mapped to questions (e.g.Q0-1), which means 2D material Interaction (Q0) and use of prediction (1) tasks to answer the question (Q0).

Query ID	**Research Questions**	**Prediction (1)**	**Classification (2)**	**Clustering (3)**	**Deep learning (4)**	**Artificial Intelligence (5)**
Q0	To what extent Image Driven ML tools have been applied to study 2D Material development and analysis?	Q0-1	Q0-2	Q0-3	Q0-4	Q0-5
Q1	To what extent Image Driven ML tools have been applied to study 2D Material engineering?	Q1-1	Q1-2	Q1-3	Q1-4	Q1-5
Q2	To what extent Image Driven ML tools have been applied to study 2D Material functional properties measurement?	Q2-1	Q2-2	Q2-3	Q2-4	Q2-5
Q3	To what extent Image Driven ML tools have been applied to study 2D Material characterization and quality assessment?	Q3-1	Q3-2	Q3-3	Q3-4	Q3-5
Q4	To what extent Image Driven ML tools have been applied to study 2D new material synthesis and engineering?	Q4-1	Q4-2	Q4-3	Q4-4	Q4-5
Q5	To what extent Image Driven ML tools have been applied to study 2D material new functional discovery?	Q5-1	Q5-2	Q5-3	Q5-4	Q5-5
Q6	To what extent Image Driven ML tools have been applied to study 2D material for corrosion application?	Q6-1	Q6-2	Q6-3	Q6-4	Q6-5
Q7	To what extent Image Driven ML tools have been applied to study 2D material for defect characterization and grain boundary?	Q7-1	Q7-2	Q7-3	Q7-4	Q7-5

2.4 Data extraction and synthesis

After the raw data (in the form of articles) were collected from the considered databases, the following information was extracted from each article: (a) the study objective (b) the findings summary (c) the source and full reference (d) the 2D material applications of interest in the study (e) artificial intelligence (AI), and the machine learning tools used to address the objective, and (f) the data science tasks involved. Furthermore, these annotations were tabulated and used to perform data synthesis to elucidate trends in the research landscape employing artificial intelligence (AI) and machine learning (ML) techniques in 2D material applications studies. The raw datasets from our paper readings were captured using Google Forms and preprocessed and analyzed using Microsoft Excel 2020 (Supplementary file S1: 10.6084/m9.figshare.12053637).

Table 2 Query formulation and search result count per database. Research questions denoted Q0 to Q7 and tasks denoted 1 to 5 in table 1 were formulated into searchable formatted queries and searched against Web of Science (WoS) PubMed, and Dimension to retrieve the papers of interest. The results are recorded in this table, for example, Q1-4 means question denoted Q1, and task denoted 4 combined. In this example, the search resulted in 04 WoS, 20 PubMed and 0 Dimension papers retrieved.

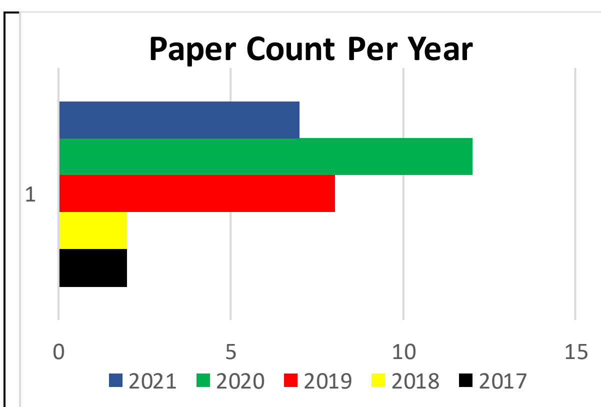
Query ID	Query	WoS	PubMed	Dimension
Q0	((optical microscope images) OR (microscopy) OR (images) or (image)) **AND** ((2D material) OR (2-dimensional material) OR (two-dimensional material)) **AND** ((machine learning))	108	152	41
Q1	((optical microscope images) OR (microscopy) OR (images) or (image)) **AND** ((2D material) OR (2-dimensional material) OR (two-dimensional material) OR (Graphene) OR (hBN) OR (MoS2) OR (WTe2)) **AND** ((Engineering)) **AND** ((machine learning))	7	61	2
Q0-1	((optical microscope images) OR (microscopy) OR (images) or (image)) **AND** ((2D material) OR (2-dimensional material) OR (two-dimensional material) OR (Graphene) OR (hBN) OR (MoS2) OR (WTe2) OR (transition metal chalcogenides) OR (TMDC) OR (Mxenes)) **AND** ((Prediction)) **AND** ((Machine Learning))	24	57	7
Q0-2	((optical microscope images) OR (microscopy) OR (images) or (image)) **AND** ((2D material) OR (2-dimensional material) OR (two-dimensional material) OR (Graphene) OR (hBN) OR (MoS2) OR (WTe2)) **AND** ((Classification)) **AND** ((Machine Learning))	28	28	7
Q0-3	((optical microscope images) OR (microscopy) OR (images) or (image)) **AND** ((2D material) OR (2-dimensional material) OR (two-dimensional material) OR (Graphene) OR (hBN) OR (MoS2) OR (WTe2)) **AND** **(**Clustering) **AND** ((Machine Learning))	7	13	2
Q0-4	((optical microscope images) OR (microscopy) OR (images) or (image)) **AND** ((2D material) OR (2-dimensional material) OR (two-dimensional material) OR (Graphene) OR (hBN) OR (MoS2) OR (WTe2)) **AND** ((Deep Learning)) **AND** ((Machine Learning))	28	63	10
Q0-5	((optical microscope images) OR (microscopy) OR (images) or (image)) **AND** ((2D material) OR (2-dimensional material) OR (two-dimensional material) OR (Graphene) OR (hBN) OR (MoS2) OR (WTe2)) **AND** ((Artificial Intelligence) OR (AI)) **AND** ((Machine Learning))	18	95	4
Q1-1	((optical microscope images) OR (microscopy) OR (images) or (image)) **AND** ((2D material) OR (2-dimensional material) OR (two-dimensional material) OR (Graphene) OR (hBN) OR (MoS2) OR (WTe2)) **AND** ((Engineering) OR (constructing) OR (manufacturing)) **AND** (Prediction) **AND** ((machine learning))	2	18	0
Q1-2	((optical microscope images) OR (microscopy) OR (images) or (image)) **AND** ((2D material) OR (2-dimensional material) OR (two-dimensional material) OR (Graphene) OR (hBN) OR (MoS2) OR (WTe2)) **AND** ((Engineering) OR (constructing) OR (manufacturing)) **AND** (Classification) **AND** ((machine learning))	7	14	1
Q1-3	((optical microscope images) OR (microscopy) OR (images) or (image)) **AND** ((2D material) OR (2-dimensional material) OR (two-dimensional material) OR (Graphene) OR (hBN) OR (MoS2) OR (WTe2)) **AND** ((Engineering) OR (constructing) OR (manufacturing)) **AND** (Clustering) **AND** ((machine learning))	0	6	0
Q1-4	((optical microscope images) OR (microscopy) OR (images) or (image)) **AND** ((2D material) OR (2-dimensional material) OR (two-dimensional material) OR (Graphene) OR (hBN) OR (MoS2) OR (WTe2)) **AND** ((Engineering) OR (constructing) OR (manufacturing)) AND (Deep Learning) **AND** ((machine learning))	4	20	0
Q1-5	((optical microscope images) OR (microscopy) OR (images) or (image)) **AND** ((2D material) OR (2-dimensional material) OR (two-dimensional material) OR (Graphene) OR (hBN) OR (MoS2) OR (WTe2)) **AND** ((Engineering) OR (constructing) OR (manufacturing)) **AND** ((Artificial Intelligence) OR (AI)) **AND** ((machine learning))	1	36	0
Q2	((optical microscope images) OR (microscopy) OR (images) or (image)) **AND** ((2D material) OR (2 dimensional material) OR (two dimensional material) OR (Graphene) OR (hBN) OR (MoS2) OR (WTe2)) **AND** **((**functional properties measurement) OR (Acoustical) OR (Atomic) OR (Chemical) OR (Electrical) OR (Magnetic) OR (Manufacturing) OR (Mechanical) OR (Optical) OR (Radiological) OR (Thermal) OR (Strain) OR (Grain boundary) OR (Conductivity) OR (Corrosion) OR (Resistance) OR (Density) OR (Ductility) OR (Malleability) OR (Elasticity) OR (Stiffness) OR (Fracture Toughness) OR (Hardness) OR (Plasticity) OR	62	129	21

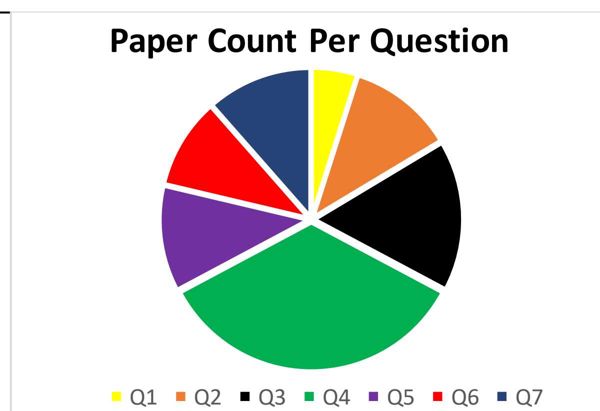
	(Strength Fatigue) OR (Strength Shear) OR (Strength Tensile) OR (Strength Yield) OR (Toughness) OR (Wear Resistance)) **AND** ((machine learning))			
Q3	((optical microscope images) OR (microscopy) OR (images) or (image)) **AND** ((2D material) OR (2-dimensional material) OR (two-dimensional material) OR (Graphene) OR (hBN) OR (MoS2) OR (WTe2)) **AND** ((characterization) OR (quality assessment)) AND (Machine Learning)	15	26	4
Q4	((optical microscope images) OR (microscopy) OR (images) or (image)) **AND** ((2D material) OR (2-dimensional material) OR (two-dimensional material) OR (Graphene) OR (hBN) OR (MoS2) OR (WTe2)) **AND** ((Synthesis) OR (Engineering) OR (constructing) OR (manufacturing)) AND (Machine Learning)	24	75	5
Q5	((optical microscope images) OR (microscopy) OR (images) or (image)) **AND** ((2D material) OR (2-dimensional material) OR (two-dimensional material) OR (Graphene) OR (hBN) OR (MoS2) OR (WTe2)) **AND** ((function discovery) OR (functional discovery) OR (discovery)) **AND** (Machine Learning)	8	8	0
Q6	((optical microscope images) OR (microscopy) OR (images) or (image)) **AND** ((2D material) OR (2-dimensional material) OR (two-dimensional material) OR (Graphene) OR (hBN) OR (MoS2) OR (WTe2)) **AND** ((corrosion) OR (corrosion application)) **AND** (Machine Learning)	1	0	0
Q7	((optical microscope images) OR (microscopy) OR (images) or (image)) **AND** ((2D material) OR (2-dimensional material) OR (two-dimensional material) OR (Graphene) OR (hBN) OR (MoS2) OR (WTe2)) **AND** ((defect characterization) OR (grain boundary)) **AND** (Machine Learning)	6	2	2
Total per database		**350**	**803**	**106**
Grand total			**1259**	

3. Results

3.1 Summary statistics

The summary statistics is a breakdown of the numbers of articles acquired from our searches of Web of Science (WoS), PubMed and Dimension databases. The PRISMA flow chart (Figure 3) illustrates the procedure in which an article or publication was excluded or included in this review. The first initial search generated 350, 803, and 106 articles from Web of Science, PubMed, and Dimension databases respectively (Table 1). In addition to these articles, 1 article was suggested and obtained from other sources of publication (home journals) during the review process. After blinded screening and eligibility curation, only 31 articles were included in this review (Figure 4, Table 3). As figure 4a shows, the distribution of articles increases from past to present (2017 to 2020), which is the indication of an increasing appreciation of AI and ML in the field of 2D materials applications and future developments. Existing literature addressed our researched questions (Table 2) with varying degree such as Q4 (new material synthesis) 24%, Q3 (characterization and quality assessment) 16%, Q2 (functional properties measurement), Q5 (new functional discovery), Q7 (defect characterization and grain boundary) 11%, Q6 (corrosion application) 10%, and Q1 (material engineering) 5%. Q4 and Q3 has significant appearance while Q2, Q5, and Q7 have reasonable presence, but Q1 is lagged in the existing literature. However, all 7 questions (Q1 – Q7) were covered to some extent, showing a diverse use of AI and ML appreciation in the domain of 2D material development and applications.





(a)	(b)

Figure 4 Paper count per year showing (a) a trend increase in AI & ML application in the study of 2D Material applications and (b) distribution across research questions and applications of AI & ML.

Table 3 Overview of key papers involved in the research questions. The 31 papers are listed with their PubMed (DOI) ID, a short description of the paper objective or goal (s), first authors' name, ML methods and ML tasks, research questions in Table 1, 2D material name, key features if available, validations method, and method accuracy.

Ref#: PubMed/DOI	Objective	Author	Year	ML Methods	ML Taks	Research Problem	2D Material Name	Features (variables list)	Validation Method	Accuracy
[35] 33707609	MS and DBSC techniques are used to segments optical microscopy images of a variety of 2D materials and substrates for a relatively high pixel accuracy from small training sets.	Sterbentz, R. M., et al	2021	Mean shift (MS), Density-Based Spatial Clustering (DBSC)	Classification, Clustering	Q1, Q3, Q4	Graphene	RGB	Probability confidence level	Pixel accuracy: 95
[10] 10.1038/s41699-020-0137-z	Deep-learning-based image segmentation, algorithm is developed and implemented for automated search, which is trained on annotated optical microscope images of 2D materials.	Masubuchi, S., et al	2020	Mask-RCNN	Prediction, Classification	Q1, Q4	Graphene, hBN, MoS_2, and WTe_2	ResNet101: position-aware high-dimensional features vectors as a JSON.	Confusion matrix	Precision ~0.53, Recall ~0.93
[36] 10.1038/s41524-019-0262-4	U-net based deep-neural network is used to identify the thickness of 2D crystals and discovered an AI-based quick exploration method for manufacturing 2D materials in large scale.	Saito, Y., et al	2019	Deep neural network (DNN)	Prediction, Classification, Deep Learning/Neuronal Net	Q4, Q5	MoS_2, Graphene	U-Net encoder extracts feature map and decode it to original size	3-fold cross-validation	70-80
[37] 10.1038/s41699-018-0084-0	A data-driven clustering analysis method is developed to automatically identify the position, shape, and thickness of graphene flakes from optical microscope images of exfoliated graphene on an SiO_2/Si substrate.	Masubuchi, S., et al	2019	Bayesian (Non-parametric mixture model)	Classification, Clustering	Q4	Graphene	Morphology features, Optical features, List (in the link)	Confusion matrix	>95
[38] 32519397	A neural-network-based algorithm is demonstrated for thickness identification of 2D materials with high prediction accuracy and real-time processing capability.	Bingnan Han, et al	2020	Deep learning, 2D material optical identification neural network (DMOINet)	Prediction, Classification, Deep Learning/Neuronal Net	Q3, Q4	Graphene	Contrast/Color, Edge, Shape, Flake Size	Confusion matrix	Table link attached
[39] 10.1016/j.eml.2020.100771	A cost-effective approach to characterizing the strength of 2D materials by processing optical microscope images of the mechanically exfoliated 2D material.	Juntan Yang, et al	2020	SVM	Prediction, Classification	Q2, Q3	Graphene	RGB	Confusion matrix	98.8
[40] 33101862	Unlike conventional optical microscopy, a contrast based on anisotropic refractive index is used to capture localized thickness-specific information for 2D material.	Abedin M. J., et al	2020	K-means (Unsupervised statistical learning method)	Classification, Clustering	Q4, Q5	Graphene / Graphene oxide	Brightness and retardance		
[41] 31976364	A neural network model was developed to classify crystal structures using small numbers of electron images and diffraction patterns with no preferred orientation.	J. A. Aguiar et. al.	2019	Naïve Bayes, Decision Forest, SVM, CNN	Prediction, Classification, Deep Learning/Neuronal Net	Q2, Q4	Graphene	Peak intensity	Confusion matrix	70-95
[42] 33566049	Lateral sizes of the exfoliated transition-metal–oxide nanosheets were predicted and controlled by the assistance of machine learning.	Mizuguchi R., et al	2021	Sparse modeling	Prediction	Q3, Q5	Exfoliated transition-metal–oxid	Table 1: - List of the explanatory variables		

[43] 10.1002/adts.202000084	A sparse modeling based prediction technique is presented as a new exfoliation strategy to generate nanosheets from soft layered composites.	Noda, K., et al	2020	Sparse modeling	Prediction	Q3, Q5, Q6	Titanium oxide: TiO	Objective and explanatory	relationship between the estimated (y') and actual (y) yields	ratio
[44] 33205523	A fully light-modulated 2D semiconductor in a simple reconfigurable phototransistor structure is presented which as a standalone device exhibits inherent characteristics of neuromorphic image pre-processing and recognition.	Ahmed, T., et al	2021	ANN	Prediction, Classification, Artificial Intelligence	Q4, Q5	Black Phosphorus	Optoelectronic (persistent photocurrent, slow recovery, long retention time)		90
[45] 33902953	A computer vision based nanopore-detection algorithm is presented which is using pixelwise precision in TEM images of nanopores in 2D membranes (WS_2).	Chen, J., et al	2021	Computer vision nanopore-detection algorithm	Prediction, Artificial Intelligence	Q2, Q3, Q4, Q7	2D membranes (WS2)	grid, radius, threshold, area	Confusion matrix	96
[46] 32932246	To disentangle the atomic distortions for two different graphene sublattices, a combination of a Gaussian mixture model enables deterministic motion of individual Si atoms in graphene along predefined trajectories.	Maxim, Z., et al	2020	Gaussian mixture model, and Principal component analysis, deep learning.	Classification, Deep Learning/Neuronal Net, Artificial Intelligence	Q2, Q5, Q6	Graphene			
[31] 32339630	Predictive and diagnostics sensing devices are designed by integrating AI and ML that can capture nanofibers, imparting high surface area, facile production, morphology control, and synergistic properties attainable.	Parshuram, MP., et al	2020	Machine Learning - AI	Prediction	Q5, Q7	Graphene			
[47] 32643230	Optimizing the substrate to achieve the maximum contrast can further extend the application of the optical microscopy method for quality control of the mass-produced graphene.	Vaziri, MRR.	2020	Mathematical optimization method	Prediction	Q3, Q4	Graphene			
[48] 32752658	Different image representations of two planar atomistic structures can identify the structures from no prior knowledge while interacting with an electronic structure program.	Christiansen, M. P. V., et al	2020	Atomistic Structure Learning Algorithm (ASLA)	Prediction	Q4	Graphene	Angular information		
[49] 10.1021/acsanm.0c02018	A new powerful (Gr-ResQ: graphene recipes for synthesis of high-quality material and pronounced as graphene rescue) is proposed, which can used as a crowd-sourced database of CVD synthesis recipes and associated experimental results to advance synthesis.	Schiller, J. A.,	2020	Template matching algorithm	Prediction, Classification	Q2, Q4, Q5	Graphen	temperatures, pressures, flows, etc. Parameters: ≈ 300		
[32] 33207552	Nonlinear optics (NLO) spectroscopy has been widely used in characterizing the physical and chemical properties of two-dimensional (2D) materials.	Zhou, L., et al	2020	SVM and more ML algorithms	Classification, Artificial Intelligence	Q2, Q4	Graphene	--	--	Max. 96.6
[50] 32132692	An image sensor constitutes an ANN that can simultaneously sense and process optical images without latency which not only train the sensor but also used to classify and encode images.	Mennel, L., et al	2020	ANN	Prediction, Classification	Q1, Q4	Graphene	36 analogue outputs $\approx 7 \times 10^{10}$ features	Train/Test Split	100
[33] 10.1002/inf2.12026	Defect topology and density in two-dimensional (2D) materials is proved to be a powerful means to tune a wide range of properties for future applications.	Dan, J., et al	2019	K-Means, Deep learning	Prediction, Classification, Deep Learning/Neuronal Net	Q2, Q4	Graphene	--	--	--
[12] 10.1016/j.jmat.2019.03.003	A combination of Fresnel law and machine learning method are useful to identify the layer counts of 2D materials.	Li, Y., et al	2019	AI (K-mean and KNN)	Prediction, Clustering	Q4	Graphene			
[51] 10.1088/2053-1583/ab1b9f	The formation of ultra-stiff diamene is exclusively found in 1-layer plus buffer layer epitaxial graphene on silicon carbide (SiC) and that an ultra-stiff phase is not observed in neither thicker epitaxial graphene (2-layer or more) nor exfoliated	Cellini, F., et al	2019	Spectral clustering	Prediction, Clustering	Q2, Q5	Graphene			

	graphene films of any thickness on silicon oxide (SiO_2).									
[52] 10.1002/adfm.201904480	Atomic defect behavior in electron beam-induced processes can be explored using a combination of deep neural networks, multivariate statistics, and Markov analysis.	Ziatdinov, M., et al	2019	DNN, multivariate statistics, and Markov analysis	Prediction	Q4	Graphene			
[14] 10.1007/s12274-018-2155-0	The machine-learning optical identification (MOI) method endows optical microscopy with intelligent insight into the characteristic color information of 2D nanostructures in the optical photograph.	Lin, X., et al	2018	SVM	Prediction	Q4	Graphene			96.78
[13] 10.1021/acsnano.8b02844	Combining genetic algorithms (GA) with MD can capture the extended structure of point defects, their dynamical evolution, and their role in inducing the phase transition between the semiconducting (2H) and metallic (1T) phase in monolayer MoS_2.	Patra. T. K., et al	2018	GA-ML	Prediction, Artificial Intelligence	Q4, Q6	Graphene			
[53] 10.1021/acsnano.7b07504	A "weakly supervised" approach is used to capture information on the coordinates of all atomic species in the image, extracted via a deep neural network, to identify a rich variety of defects that are not part of an initial training set.	Ziatdinov, M., et al	2017	Deep learning	Deep Learning/Neuronal Net	Q3, Q6	Graphene			
[54] 27867837	A definition of a structural state that is composed of both local and nonlocal information extracted from atomically resolved images and is wholly untethered from the familiar concepts of symmetry and periodicity.	Laanait, N., et al	2017	LoG (detector in the Python scikit-image library [41] and SIFT in OpenCV [42])	Prediction, Classification	Q3, Q4	Graphene			
[55] 10.1039/d0na00634c	A transfer learning approach is used to recognize and classify airborne CNT/CNF particles from TEM images by using hyper column vectors, which were clustered via K-means and processed into a Vector of Locally Aggregated Descriptors (VLAD) representation to train a SoftMax classifier.	Luo, Q., et al	2021	K-means, Gradient boosting algorithm	Classification, Clustering	Q4, Q6	Graphene	1472 features per pixel		90.9, 84.5
[56] 30968576	Predicting structure-property relations and reduce dependence on simulations by using low dimensional physical descriptors to statistically describe the defects, which shows the purpose-specific microstructure representation at low computational cost.	Hundi, P., et al	2019	Deep Learning	Prediction, Deep Learning/Neuronal Net	Q2, Q7	hBN and Graphene	RGB	Train/Test Split	>95
[57] 34106715	Structural engineering of low-dimensional structures using a near ultrahigh vacuum system comprised of an aberration-corrected scanning transmission electron microscope is useful for large sample areas.	Trentino, A., at al	2021	CNN	Deep Learning/Neuronal Net	Q3, Q7	Graphene	RGB		
[58]10.1088/2053-1583/abd72c	Deep neural network (DNN) incorporated into a motorized microscope that automatically scans entire silicon wafers to detect and identify two-dimensional (2D) materials.	Shin, Y. J., et al	2021	GT-DNN	Prediction, Classification	Q3	Graphene, hBN	red, blue, cyan and yellow	Confusion matrix	91.9

3.2 Q0 – Current use of image-driven ML techniques to solve issues involved 2D materials

Advances in 2D materials and exponential growth in their applications pose multifaceted research challenges. These challenges can be addressed with improved interdisciplinary knowledge, developing new diverse technologies, and techniques. Data (spectroscopy, microscopy, text, etc.) driven computing paradigm is one such technique which is significantly contributing to solve challenging issues involved 2D materials. For example, challenges like number of layers, crystal orientation, crystal phase, chemical specificity, strain, chemical dynamics, etc. are addressed using nonlinear optics (NLO) spectroscopy data, and defect topology [32] is analyzed using data from scanning transmission electron microscopy (STEM) [33]. In both the cases data driven computing paradigm is adopted to extract meaningful knowledge [32], [33]. This review is focusing on current research work answering questions involved 2D materials using microscopic images and spectroscopic data driven artificial intelligence (AI) and ML techniques. The goal is to assess up to what extent image driven AI and ML techniques can solve issues involved 2D materials and answer fundamental research queries. In addition, this review may draw attention towards research gap, open questions, and shortcomings in the current research work and provide future research directions. In subsequent sections, issues such as material engineering, functional properties measurement, characterization and quality assessment, new material synthesis and discovery, corrosion applications, defect characterization and grain boundary will be discussed with focus on "*microscopic images and spectroscopic data*", "*2D materials*", and "*AI & ML techniques*".

3.2.1 2D material thickness and layers (Q1 – Q4)

Developing automated, cost-effective, and efficient tools/techniques to accurately quantify thickness, and determined number of layers in a 2D wafer is a daunting task. This task required multidisciplinary knowledge, efficient algorithms, computing resources, and significantly large data sets for better accuracy of results. Diverse range of imaging data (e.g., optical microscopy, spectroscopy) are used to measure functional properties, assess quality, and synthesize new 2D materials using computational power of artificial intelligence (AI) and machine learning (ML) techniques. Measuring strengths of 2D materials not only used to understand their functional properties but also to quantify qualitative assessment by using mechanical, chemical, electrical, structural characterization (Q2–Q3), material engineering (Q1), and synthesizing the new materials (Q4). Here, we briefly discuss the Q1 – Q4, related issues and proposed solution methodologies. Machine Learning (ML) techniques are broadly classified as supervised and unsupervised ML algorithms; therefore, first we explore queries Q1 – Q4 thar were being solved using supervised ML techniques.

Two-dimensional (2D) material thickness identification (or prediction) can be discovered using optical microscopic images (OMI) with computational power of supervised ML techniques. An automated OMIs are segmented using supervised machine learning algorithm (i.e., deep learning) to detect exfoliated 2D materials on SiO_2/Si substrates. OMIs are segmented to automatically scans entire silicon wafers to detect and identify 2D materials. Different sizes, shapes, and thicknesses of graphene was classified with approx. 91.9% accuracy [58], and with a recall of ~0.93 [10]. The proposed algorithms were successfully detecting number of layers and thickness of 2D materials. In addition, color features are instrumental to predict the most relevant physical properties of 2D materials. In [14], segmented OMIs gave useful insight through color information of 2D nanostructures on a large-area of graphene to identify thickness, impurities, and stacking order (of layers). Mechanical exfoliation is a useful method to manufacture 2D crystals for various practical applications. In [43], OMIs were used as input data to identify thickness with deep graphical features such as contrast, color, edges, shapes, flake sizes, and their physical distributions. These 2D crystals contain relatively thick flakes on substrates, which need to be removed for efficient manufacturing of atomic 2D crystals. OMIs segmentation was used to identify the thickness of 2D atomic layer flakes in [36]. OMIs are taken as input to quantify the statistical distribution of 2D material flakes on their different layers. The statistical distribution of the flakes' size is used to estimate the strength of the associated 2D material (i.e., graphene) [39], which is useful for various strength-sensitive applications. In [59], a big data analysis and deep learning methods is developed as a set of visible-light hyperspectral imaging technologies, which is can identify few-layers of MoS_2.

Unlike supervised ML, unsupervised ML draw inferences and find patterns from input data without referring to the labels of the data. Now, we briefly discuss about Q1 – Q4 with solution using unsupervised ML techniques in the existing work.

Identifying (or predicting) thickness of 2D materials can also be achieved using optical microscopic images (OMIs) with computational power of unsupervised ML techniques. Clustering algorithm grouped the OMI data in red, green and blue (RGB) space to determine/identify thickness of the 2D material layers with a pixel accuracy of 95% [35]. Clustering analysis method is used to automatically identify the position, shape, and thickness of graphene flakes from OMI of exfoliated graphene. This analytical method can be applied to a range of substrates with differing SiO_2 thicknesses [37]. Thickness-specific information can also be extracted by using birefringence images of graphene dispersions with contrast based anisotropic refractive index in 2D materials. The proposed clustering algorithm was extracting pixel-by-pixel information from brightfield and birefringence images, and clusters show nanoplatelets (partially exfoliated), and 2D sheets (well-exfoliated) species [40]. In [51], a combination of atomic force microscopy (AFM) topography and friction force microscopy (FFM) images are used to identify thickness of graphene domains in epitaxial and exfoliated films**.** A clustering algorithm provide a new finding of a few-layer thick epitaxial graphene into a new ultra-hard carbon phase (named diamene) on the pressure-induced chemical transformation.

Unlike individually using supervised and unsupervised ML algorithms, both techniques can also be used simultaneously. For example, an automatic image sensor is engineered to classify (supervised ML) and encode (unsupervised ML). This reconfigurable 2D semiconductor was using artificial neural network (ANN) to train the sensor which simultaneously sense and process OMIs. The ANN used by autoencoder can learn an efficient representation (encoding) for a set of images, and input-out pair for the ANN training data was generated from the unsupervised ML [50]. Similarly in [38], a semantic segmentation network based encoder-decoder is presented for pixel-wise identification of OMIs for 2D materials. This encoder-decoder can identify various 2D materials in OMIs regardless of variations in optical setups in real time. In [47], two analytical equations are derived to find the accurate thickness of the SiO_2 layer in graphene/SiO_2/Si structures. These substrates are in common use for fabrication of graphene-based devices.

3.2.2 2D material defects, grain boundary and crystallographic structure (Q5 – Q7)

The latest issues involved 2D materials such as functional discovery (Q5), corrosion applications, defects detection (Q6) and grain boundary (Q7) (Table 2). Various types of input data like optical microscopic images (OMI), transmission electron microscopy (TEM), scanning transmission electron microscopy (STEM), aberration-corrected scanning transmission electron microscope (ACSTEM), high-resolution transmission electron microscopy (HRTEM), Raman spectra (RS), and dynamic light scattering (DLS) are used with computational power of supervised ML techniques to study 2D materials strengths, defects, and crystallographic structure to redesign existing and develop new applications.

Defects:

Controlling, assessing, and identifying defects is an important and challenging task due to inherent delicate structure of 2D material. Atomic scale imaging through transmission electron microscopy (TEM) leverage to develop 2D materials fabrication tool. These tools help to extract deep knowledge from the images which can be used to detect 2D nano porous, monolayer membranes, and assess defects with grain boundaries as a source of measurements. Such tools has several applications like ion detection (or conductance), water filtration, and DNA sequencing [45]. In [57], aberration-corrected scanning transmission electron microscope (ACSTEM) data is used to develop atomically clean and free-standing graphene with a controlled defect distribution. The defects information provides statistical knowledge on structural changes of the S_i atom neighborhood during atomic motion. Capturing defects by manipulating electron beam in a deterministic motion of individual S_i atoms in graphene leverages controlled atom-by-atom fabrication [46]. Oxidation-related defects in 2D black phosphorus (BP) is exploited to achieve visual memory, wavelength-selective multibit programming, and erasing functions. Neuromorphic information based computational methods are used to classify these types of defects [44].

Extracting local structural information is useful in defect-detection on a multilayer graphene surface. The structural information is extracted using TEM images in the presence of common factors such as noise, limited spatial resolution, and weak contrast [54]. Defects with angstrom-level precision can be identified using atomic scale images. The goal is to interpret complex atomic scale defect transformation, including switching between different coordination of silicon dopants in graphene as a function of time, etc. [53]. Structural defects govern various physical, chemical, and optoelectronic properties of two-dimensional transition-metal dichalcogenides (TMDs). Methods using high-resolution transmission electron microscopy (HRTEM) data can interpret point defects in structure, their dynamical evolution, and their role in inducing the phase transition between the semiconducting (2H) and metallic (1T) phase in monolayer MoS_2. This work significantly advances the current understanding of defect evolution at structural level and structural transitions in 2D TMDs, which is crucial for designing nanoscale devices with desired functionalities [13].

Structural properties:

Structural properties of the 2D material depends critically on chemical physics. Two planar atomistic structures (ideal graphene and graphene with a grain boundary region) influence the ability to identify the structures from no prior knowledge. Image driven angular information is also useful to delicately analyze structural properties [48]. E-beam irradiated S_i atoms in the bulk and at the edges of single-layer graphene is examined using scanning transmission electron microscopy (STEM). A deep learning network is used to convert experimental STEM movies into coordinates of individual S_i and carbon atoms. The proposed model used to establish the elementary atomic configurations of the S_i atoms, defining the bonding geometries and chemical species and accounting for the discrete rotational symmetry of the host lattice. This analysis enables insight into the defect populations and chemical transformation networks from the atomically resolved STEM data [52]. Similarly, extracting crystallographic structural information can further help to deeply understand and synthesize 2D materials. This structural information is extracted from atomic resolution image or diffraction patterns using supervised ML algorithms [41].

Importance of measuring the structural properties of 2D materials specifically in nanoscience is not only significant for the quality assessment and applications but also for the functional discovery. A computational method to recognize and classify airborne carbon nanotubes/nanofibers (CNTs/CNFs) particles from TEM images. The proposed model automatically detects and classify complex carbon nanostructures with potential applications that extend to the automated structural classification for other nanomaterials [55]. Controlling lateral-size of the 2D materials is caused by the breaking-down processes. The lateral size of the nanosheets was estimated by dynamic light scattering (DLS) to achieve better analysis. The eight physicochemical parameters of the organic guests and dispersion media were extracted by sparse modeling for the construction of the size-prediction model. The size-prediction model accelerated the selective syntheses of the nanosheets with large and small lateral sizes [42].

Synthesis:

The graphene integration and fabrication process during manufacturing electronic devices requires sustainable production of graphene on dielectric substrates. Viable industrial manufacturing of graphene should ideally conserve the catalyst and reaction gases, but typically used metal catalysts that are dissolved after synthesis. This is due to hinderance caused by the hundreds of coupled synthesis parameters. These sensitive parameters may strongly affect chemical vapor deposition (CVD). To address this issue, a data-driven graphene synthesis tool Gr-ResQ (i.e., graphene rescue) is developed, which carefully selects parameters using Raman spectra and microscopy images data. Later, these parameters advance the new graphene synthesis process using computing power of machine-learning algorithms [49]. In [12], a layer count for 2D materials on Si/SiO_2 substrate is presented using red-green-blue index. The Wafer-scale synthesis is executed on a large-area 2D nanostructures. The proposed method is proved to be comparatively fast and accurate.

Functional discovery:

Contributions of component materials used in sensing have significant impact on the utility of nanofibers in various applications like oximetry, volatilome detection, and biomarker detection. Fluid sample sensing and imaging by processes like surface deposition on nanofibers, immobilizing, calcination, etc. have the strong influences on sensing device properties. Functional graphene foam with porosity and functional group presence was forming a strong covalent bond resulting in stability [31].

3.3 Pictorial representation of image driven tasks, AI & ML techniques, and 2D materials

Three parts of the figure 5 illustrate the focus of this review using word cloud from the literature tabulated in the Table 2: (a) AI & ML techniques, (b) tasks, and (c) 2D materials & substrates. Graphene is the most used term to represent 2D materials with hBN and MoS_2 substrates in the collected literature as shown in figure 5c. Although, collected literature contains five different tasks such as classification, prediction, clustering, deep learning, and artificial intelligence (AI), but classification and prediction are depicted as the most required tasks shown in the figure 5b. Figure 3a shows the computational methods used to solve the issues involved 2D materials with images as input data. Convolutional neural network (CNN), deep neural networks (DNN), recurrent CNN (RCNN) are the algorithms based on the computational philosophy of artificial neural networks (ANN) to analyze visual imagery. Figure 5a shows that CNN and its variants are the advance versions of ANN with convolution kernels (as hidden layers) is the most popular computational method used in collected literature for this review. In addition to CNN, some other computational techniques like sparse modeling, Naïve Bayes, K-Means, and Artificial Intelligence (AI). Note that AI is refereed as combination of computational techniques such Genetic Algorithm and ML as in [13].

(a) AI & ML methods	(b) Tasks	(c) 2D materials & substrates

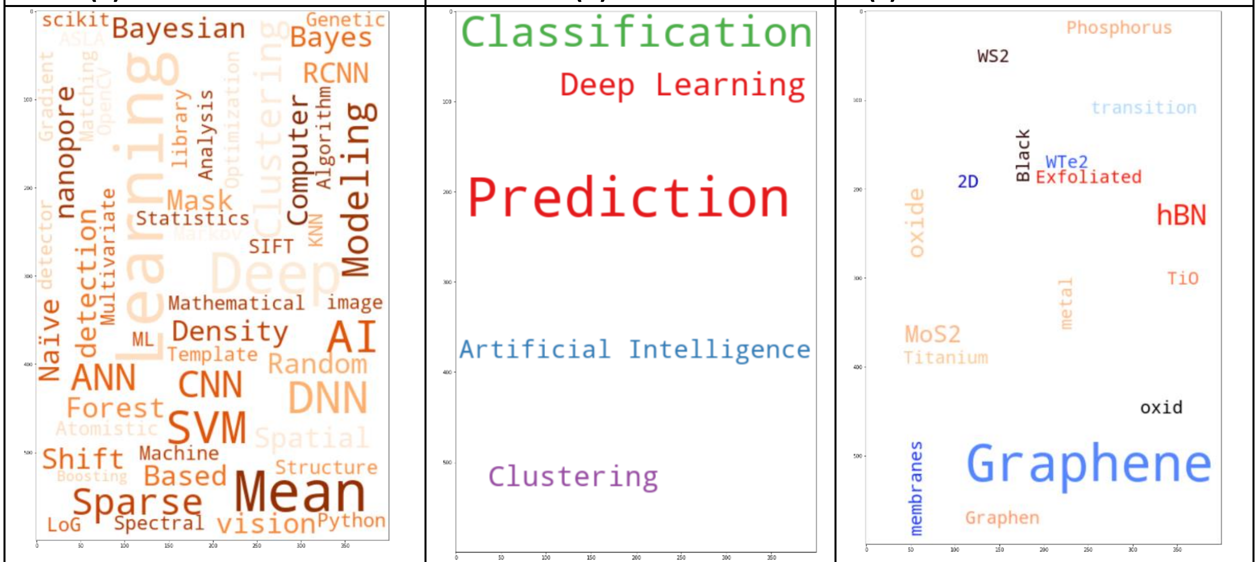
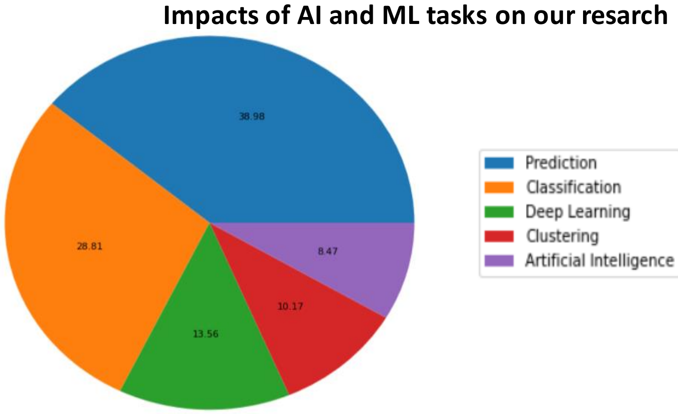
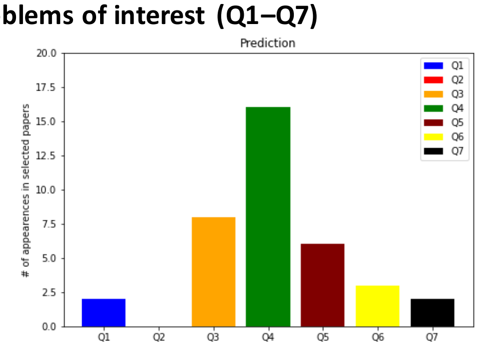


Figure 5 Pcitorial view of AI & ML methods, tasks and materials involved in the selected papers (n = 31). Fig. 5 was created with word-cloud by using (a) AI & ML methods, (b) tasks, and (c) 2D materials and substrates. The visibility of a word among words in their panel emphasize their appearance in papers, which highlights a review papers' focus.

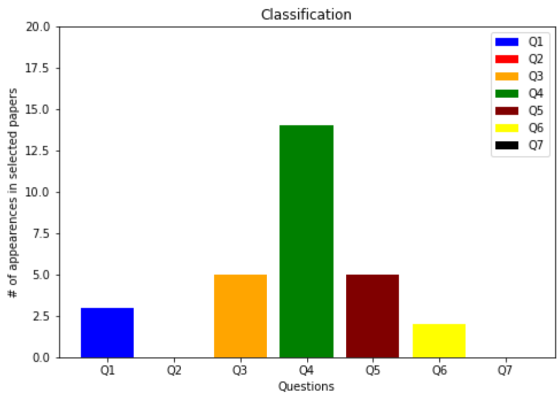
Figure 6 illustrates graphical distribution of AI & ML task versus number their appearances in the articles we collected for this review as shown in Table 2, 3. Prediction, classification, deep learning, clustering, and AI tasks appeared 38.98%, 28.81%, 13.56%, 10.17%, and 8.47% respectively as shown in pie chart in the 6. This shows that prediction is the most widely unsolved task to solve by the research community while AI is the least highlighted task. Note that the task AI refers to the combination of two or more ML methods or the combination of one ML and one other method to solve the data driven problem such as in [12] [13]. Out of 5 ML and AI tasks, the bar charts in figure 6 shows that Q4 is dominated in prediction, classification, deep learning, and clustering tasks while Q3 is dominated for AI task. Q2 is not appeared for any task while Q7 is the least appeared only for the deep learning task. Q1 is not appeared for the deep learning and AI tasks while Q3–Q5 are appeared for the AI task. We can conclude that currently the most studied issues are 2D new material synthesis and engineering while using image driven ML tools. On the other hand, work on corrosion application with "*2D materials*", "*image-driven*", "*AI and ML*" is not done yet as per our search. Unlike tools for other tasks, algorithms for the task AI are not much explored yet.

Impacts of AI and ML tasks on our resrach problems of interest (Q1–Q7)

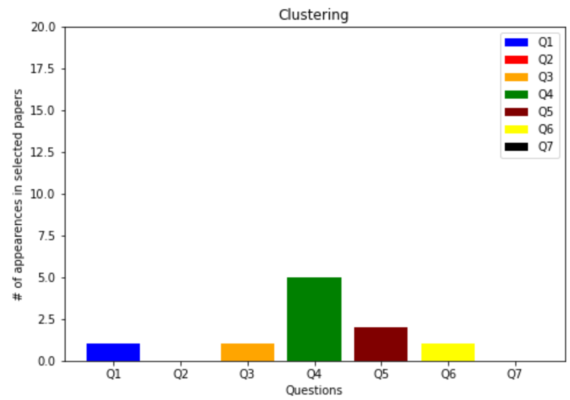




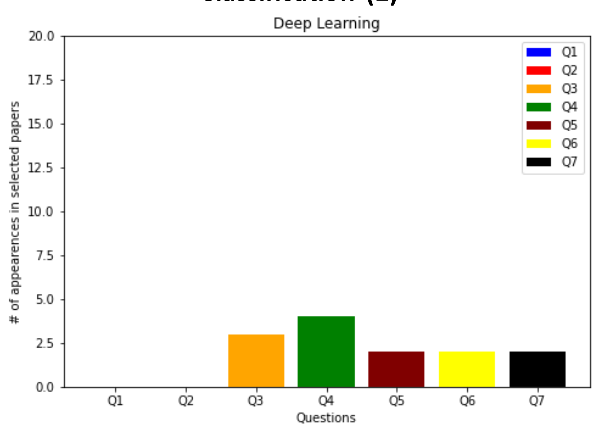
Prediction (1)



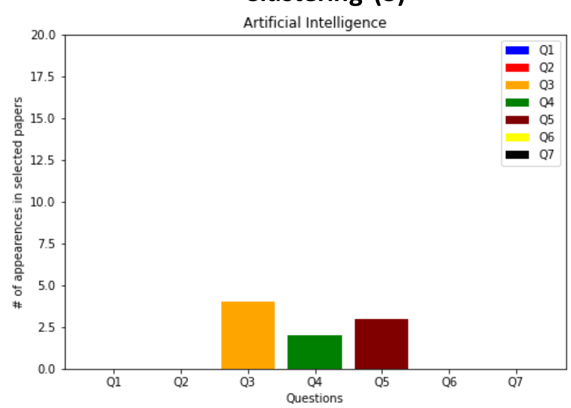
Classification (2)



Clustering (3)



Deep learning (4)



AI (5)

Figure 6 Quantitative representation of various ML tasks involved in the selected papers.

Figure 7 shows another word cloud for a broader view of knowledge and point of emphasis in the current research articles as shown in Table 3. Keywords list, short objective, and title of the included articles may provide this big picture. 2D material, graphene, types of microscopic images are the main point in articles keywords list as shown in Figure 7a. Short objective and titles of the included articles shows also show a clear emphasis 2D materials, graphene, and images. In addition, short objective shows emphasis on the thickness of 2D materials while titles show importance of exfoliation process. Overall, this word cloud also shows the relevancy of the articles to our key questions and three keywords or terms "*2D materials*", "*image-driven*", "*AI and ML*".

(a) Articles keywords	(b) Articles short objectives	(c) Articles titles

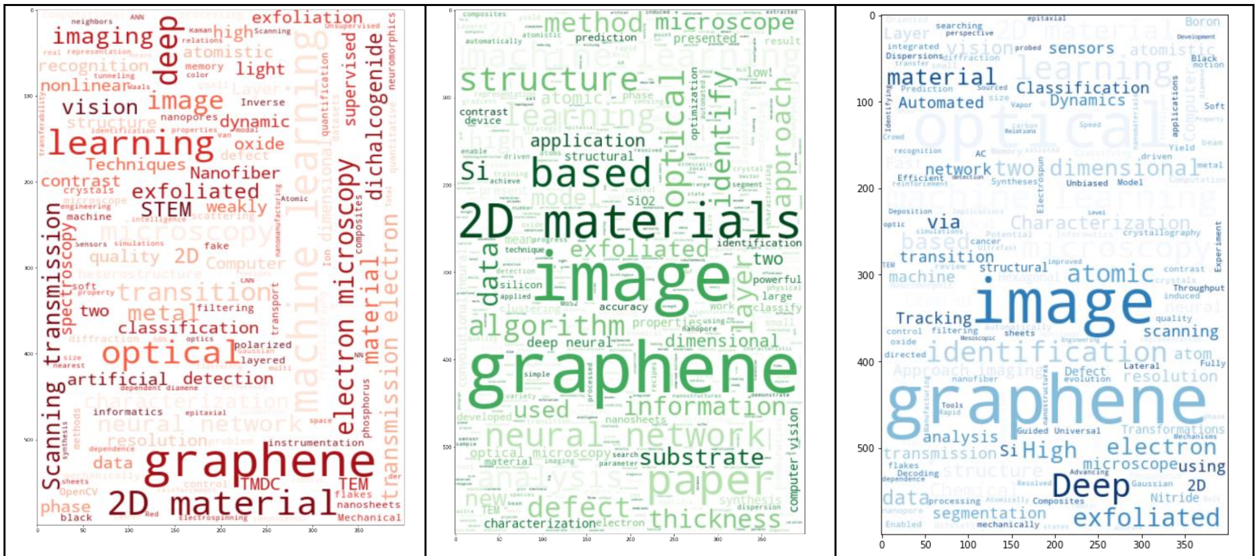


Figure 7 Emergence of important terms in the various parts of the selected papers (n = 31). Fig. 5 was created with word-cloud by using (a) papers key words, (b) words contained in study objectives (c) titles of the paper. The visibility of a word among words in their panel emphasize their appearance in a papers' keys words, objective, or title, and highlight a review papers' focus.

4. Discussion

4.1 Challenges and opportunities

Data driven computing models are becoming popular due to better data sources, and relatively easy and cheap availability of computing resources. These scenarios not only present numerous opportunities for diverse 2D material applications but also pose serious research challenges in terms of interdisciplinary knowledge and efficient and accurate use of resources. We identified some key challenges opportunities to the research community during this review process. Some of these challenges and opportunities are the data availability, ML/AI algorithm diversity, heterogenous benchmark comparisons, etc. As focus of this review is on various types of microscopic images, spectroscopic data, or combination of microscopic/spectroscopic data; therefore, pertaining to the new developments, applications, engineering, and functional discovery for 2D materials data is not publicly available in most of the articles listed in the Table 3 except some papers such as [10].

4.4.1 Data as input

The opportunities created by nanofiber research is rewarding to address the existing challenges of cost-intensive and complicated apparatus for cancer detection applications. These devices can be easy to use for practitioners at healthcare facilities with high reliability and cost-effectiveness [31]. A nonlinear optic (NLO) spectroscopy is being used to identify the 2D material structure. Currently, limited systems are only suitable for a certain NLO characterization method. With spectrally narrow excitation light for broadband spectrum detection, the excitation wavelength requires to be adjusted continuously. This continuous variation makes the processes tedious which not only inconvenient to use but also inefficient due to low characterization. This may also increase level of inaccuracy in results due to the long detection time which may hinder rapid crystal phase transition in real time [32].

4.1.2 Computational techniques

Recently, artificial intelligence (AI) and machine learning (ML) algorithms are vastly used to quantify the defect characterization, thickness, and structure identification of 2D materials. Many of AI & ML algorithms used microscopy and spectroscopic data as input. Generally, these algorithms are categorized as supervised and unsupervised ML. In supervised ML, lack of generalized training data set is hindrance in structure identification process for a broader range of materials. In other words, extensive retuning of supervised ML algorithm parameters is needed for structure identification of each 2D new material. Therefore, current implemented supervised ML algorithms are only self-consistent for a small data set. Similarly, unsupervised ML algorithms are also lacking in detailed comparative study on various unsupervised methods for diverse data sets. To this end, exploring an ideal combination of different methods for optimum results is itself a tedious task. In addition, lack in homogenous validation criteria for different representation and identification methods is also an obstacle in standardized performance comparisons for different algorithms [33].

4.1.3 Data modalities

In this review, the research articles involved 2D materials with microscopic images, spectroscopic data, or combination of microscopic/spectroscopic data in addition use AI & ML algorithms as data driven techniques. The articles included in this review (in Table 3) used various types of images and signal data. Optical microscopic image (OMI) is a commonly used image type for research areas including microbiology, microelectronics, nanophysics, pharmaceutics, biotechnology, and histopathology. Highly magnified images (up to 2 million times) can be generated by using a particle beam of electrons with transmission electron microscopy (TEM) to visualize specimens such as tissue sections, molecules, etc. While Scanning transmission electron microscopy (STEM) is working on the principles of TEM and scanning electron microscopy (SEM) to produce very finely focused beam of electrons across the sample in a raster (graphic) pattern. Aberration-corrected scanning transmission electron microscope (ACSTEM) is an extension of STEM and working on same principles which is allowing focus of images to sub-angstrom resolution level. This fine focused level identifies individual atomic columns with unprecedented clarity. Similarly, the high-resolution transmission electron microscopy (HRTEM) uses both the transmitted and the scattered beams to create an interference image which is as small as the unit cell of crystal. On the hand, atomic force microscopy (AFM) is a very-high-resolution type with demonstrated resolution on the order of fractions of a nanometer, more than 1000 times better than the optical diffraction-limit. Friction force microscopy (FFM) is same as AFM and provides information on the properties of molecular materials.

Extracting and analyzing 2D material information using spectroscopic data is also a popular technique. Spectroscopic dynamic light scattering (DLS) is a technique used the size and size distribution profiles of particles with Brownian motion. While Raman spectroscopy (RS) is a non-destructive chemical analysis technique to provide detailed information about chemical structure, phase and polymorphic, crystallinity and molecular interactions for 2D materials. RS features several peaks, showing the intensity and wavelength position of the Raman scattered light. On the other hand, a nonlinear vibrational spectroscopy (NLO) is working with two colors of light that are mixed at an interface to create the sum frequency. The generated sum frequency light carries structural information about the molecules and chemical functionalities present at the interface. The visible hyperspectral imaging technology (VIS-HSI) with combination of charge-coupled device (CCD) is used as a visible hyperspectral algorithm (VIS-HAS) in [59]. The MoS_2 sample is measured using a Raman microscope to determine the position of each layer distribution, and an image is taken using an optical microscope and CCD to capture the same position. The obtained CCD image is combined with VIS-HSI to convert the spectral characteristics of each layer of MoS_2, and data preprocessing is performed. Multiple data types are opportunities as well as challenges at the same time. Diversity and range of microscopic and spectroscopic data implies multi-level difficulty due to tradeoff: (a) types of apparatus to generate data, (b) availability of resources and domain knowledge, (c) AI and ML techniques to solve specific problems.

4.1.4 Data availability & reproducibility

In addition to the computational code (software), data availability is important step towards reproducing the results. Only few of the listed articles in the Table 3 provide the public availability of the software and datasets, which limits the reproducibility. As mentioned in [34], data reproducibility is the gold standard of the data mining and machine learning domain. Therefore, we observed that only few studies included in this review could be reproducible due to software and datasets unavailability. For the sake of model usefulness, the data used to make the model available and easily accessible to those who want to reproduce it. Arguably, the datasets used in the articles we reviewed might meet the standard, if both software and datasets availability were addressed adequately.

4.2 Future directions

4.2.1 Knowledge discovery

Leveraging microscopic and spectroscopic data to discover useful knowledge by AI and ML techniques is becoming popular due to computational resources and data availability. Capturing properties of atomically thin 2D materials is an extremely complex task and synthesizing such material is even more sensitive in nature. Similarly, growth process, thickness, crystallographic structure, defect characterization, and coatings need AI and ML computational techniques to extract computationally sensitive features for better understanding and future developments using 2D materials. Therefore, a reliable and efficient growth and manufacturing process is crucial to synthesize 2D materials at the wafer scale such as flexible and transparent optoelectronics [60]. Detecting number of layers and thickness of 2D materials while extracting color features by using microscopic images may be proved instrumental to predict the most relevant physical properties of 2D materials. These color features of 2D nanostructures not only useful to identify thickness but also existing impurities, and stacking order (of layers) for various practical applications [14]. Defect-detection in 2D material is important either it is generated due to rough conditions or introduced deliberately on a multilayer graphene surface. These defects can be identified up to angstrom-level precision using atomic scale microscopic images [53] [54]. 2D material coatings exhibit complex and sensitive interactions with biological matters. Antibacterial coatings to prevent biofouling, biocompatible platforms suitable for biomedical applications (e.g., wound healing, tissue repairing and regeneration, and novel biosensing devices) could be useful in near future. Deep understanding of 2D material coatings may modulate a specific bacterial or cellular response, which is critical for any future innovation in the field [3]. 2D material with microscopic images. To this end, microscopic and spectroscopic based data driven AI and ML techniques may play an important role in discovering the fundamental knowledge which leads to develop new applications using 2D materials.

4.2.2 Leveraging new ML/AI techniques

Artificial intelligence (AI) and machine learning (ML) techniques are useful to demystify core knowledge through various types of microscopic/spectroscopic data for 2D materials which not only leads to understand their structure, properties, and functions but also open doors for new applications. This core knowledge is the quantification of the 2D material attributes which is used to understand the structure orientation, identification, and functional properties with microscopic images are used as input [33]. As spectroscopic data, the use of nonlinear optical (NLO) signals is still in early stage of characterizing various attributes of 2D materials such as judging the position of the thinner layers. In addition, defects identification and variation, and strain direction and engineering are also important points to understand 2D materials for better future applications using NLO signals. Therefore, power of ML methods are considered as better potential candidates to implement the laws of NLO signals to achieve the goal of classifying and quantifying the characteristics of 2D materials [32].

ML algorithms can be used to correlate the structures and properties of 2D materials. For example, a combination of ML with molecular dynamics (MD) with high-resolution transmission electron microscopy (HRTEM) as input is simulated to explore the structural optimization and evolution of defects to help understand the structural transition in 2D materials. Similar experiments are performed while combining ML with density functional theory (DFT) calculations, structural and electronic properties of different hybrid 2D materials, and to screen various parameters for vdW heterostructures. A force field for classical simulations of stanene was developed using a ML method and trained by data sets from ab-initio results to calculate the mechanical and thermal properties of stanene. 2D materials without bulk layered counterparts are also being discovered, using evolutionary algorithms like genetic algorithms, particle swarm optimization, where both evolutionary algorithms are using energy-based merit criteria and additional biases towards 2D sheets [60].

4.2.3 Model selection

Supervised and unsupervised ML are two models that are commonly used to solve various issues using microscopic/spectroscopic data that involve 2D materials as listed in Table 3. For example in Table 3, Mask-RCNN as a supervised ML algorithm is used in [10], while Mean shift Density-Based Spatial clustering as unsupervised ML algorithm is used in [35]. Third way is to use combination of algorithms such as k-means and K-NN algorithms in [13], ML and genetic algorithms in [12], and mathematical optimization in [47] to solve the various issues as listed in the Table 3.

Clustering is a branch of unsupervised learning algorithms involving algorithms like K-means, hierarchical clustering, density-based spatial clustering, etc. The performance of such algorithms depends on quality of the input feature vectors. Selecting with a reasonable tradeoff from a pool of clustering algorithms with existing scenarios and requirements need with in depth knowledge about advantages and disadvantages of these algorithms. Issues like identifying atomic structures, crystal structure, defects characterization are being solved with varying level of accuracy by using the clustering algorithms [33].

The deep learning models generalize microscopic structural data and is dependent on the training data set. A reasonable amount of training microscopic data is simulated with some augmented methods. However, the construction of the simulated images for training relies on prior knowledge, which is extracted through annotation procedure manually. Deep learning techniques such as CNN, RCNN, or Mask-RCNN have sparked intense research interest in the fields of image classification, natural language processing, machine translation, and material informatics as listed in Table 3. It should be noted that deep learning can be supervised, unsupervised, semi-supervised, or reinforced. In stark contrast to unsupervised learning, supervised deep learning algorithms learn a nonlinear function that directly correlates the input atomic structure to a pretrained structure [33].

4.2.4 Performance evaluation metric and accuracy

We observed different types of performance evaluation metrics in this review such as cross-validation, confusion matrix, and train-test split methods. Out of 31 articles qualified for this review, 7 articles used confusion matrices, 2 used train-test split, and 1 article used 3-fold cross-validation and rest of the articles did not provide clear information on their performance evaluation metrics. Because of variability and heterogeneity in quality and quantity of data and computational technologies, it is also important to provide quantified accuracy in their results, which is missing at this point as shown in the Table 3. Out of 31 articles qualified for this review, 16 articles provide quantified accuracy of their computational methodologies, and rest of the articles did not provide clear information about the accuracy as listed in the Table 3. Therefore, in the future, a standardized and homogeneous performance evaluation metrics and accuracy criterion is important for 2D material research.

4.2.5 Feature engineering

Feature engineering is the process of using domain knowledge to select and translate the relevant variables from the given data for ML or statistical modeling. AI and ML algorithms heavily rely on feature engineering. Out of 31 articles qualified for this review, in most of the cases color features are extracted to process the microscopic/spectroscopic data as listed in the Table 3. These features include morphological, RGB (red, green, blue), RGB with cyan, image intensity, edge, shape, flake size, radius, area, temperature, pressure, flow, etc. Accuracy of most classifiers affected in case of using unrelated features, and too many features are computationally time consuming. Therefore, a balance by feature engineering could save a training time and resources while addressing 2D material data. Also, note that the application of feature engineering in this domain is as important as building a data driven model itself.

5. Conclusion and future work

In this review, we explored the concepts of artificial intelligence (AI) and machine learning (ML) as applied towards understanding 2D materials' crystallographic structure, functional properties, function discovery, quality assessment, defects, corrosion detection, synthesis, grain boundaries, and their implications on 2D materials' existing and upcoming applications. From the articles we reviewed, 38.98% studies involved predictive models using supervised machine learning, while 10.17% and 13.56% used unsupervised methods and deep learning respectively. In the retrieved articles, prediction (38.98%) and classification (28.81%) were among the most dominant machine learning tasks, which were used to classify and predict relevant features that dictate 2D materials' knowledge discovery (e.g., structure, properties, synthesis, etc.). Furthermore, the utility of diverse data modalities together with various methods to engineer or select image features proved valuable in many of the reviewed studies. While AI and ML are being increasingly applied in many 2D material and related domains (material design [11] , defects [33], vector-host pathogen [34]) as shown in this review, they have not yet taken roots in the field of knowledge discovery for 2D materials using microscopic images or spectroscopic data.

Corrosion detection, application, and prediction for 2D materials is still an issue to explore using AI/ML techniques while microscopic images and spectroscopic data as input. In general, clustering, deep learning and AI lagged the other tasks of ML, such as classification, and prediction of 2D materials using microscopic images or spectroscopic data. A future increase in deep learning applications in the field could be valuable, especially when combined with other approaches such as AI, cross-validation, and model selection. Also, the application of AI (i.e., combination of two approaches: ML, evolutionary computation, mathematics, etc.) would increase hypothesis generation in the field and reduce the time and resources spent in doing so. Increase in AI & ML use in the domain is still facing many challenges, these approaches have great potential and should be encouraged to bring new perspectives to old and diverse problems involved 2D materials.

6. Reference

[1] P. Gao, Y. Xiao, YuliangWang, L. Li, W. Li, and W. Tao, "Biomedical applications of 2D monoelemental materials formed by group VA and VIA: a concise review," *J. Nanobiotechnology 2021 191*, vol. 19, no. 1, pp. 1–23, Apr. 2021, doi: 10.1186/S12951-021-00825-4.

[2] G. Chilkoor *et al.*, "Hexagonal Boron Nitride: The Thinnest Insulating Barrier to Microbial Corrosion," *ACS Nano*, vol. 12, no. 3, pp. 2242–2252, Mar. 2018, doi: 10.1021/ACSNANO.7B06211.

[3] J. Santos, M. Moschetta, J. Rodrigues, P. Alpuim, and A. Capasso, "Interactions Between 2D Materials and Living Matter: A Review on Graphene and Hexagonal Boron Nitride Coatings," *Front. Bioeng. Biotechnol.*, vol. 0, p. 5, Jan. 2021, doi: 10.3389/FBIOE.2021.612669.

[4] G. Fiori *et al.*, "Electronics based on two-dimensional materials," *Nat. Nanotechnol. 2014 910*, vol. 9, no. 10, pp. 768–779, Oct. 2014, doi: 10.1038/nnano.2014.207.

[5] S. Haastrup *et al.*, "The Computational 2D Materials Database: High-throughput modeling and discovery of atomically thin crystals," *2D Mater.*, vol. 5, no. 4, 2018, doi: 10.1088/2053-1583/aacfc1.

[6] "Graphene applications." https://www.graphene.manchester.ac.uk/learn/applications/ (accessed Jul. 01, 2021).

[7] R. Ramprasad, R. Batra, G. Pilania, A. Mannodi-Kanakkithodi, and C. Kim, "Machine learning in materials informatics: Recent applications and prospects," *npj Computational Materials*, vol. 3, no. 1. Nature Publishing Group, p. 54, Dec. 01, 2017, doi: 10.1038/s41524-017-0056-5.

[8] J. Schmidt, M. R. G. Marques, S. Botti, and M. A. L. Marques, "Recent advances and applications of machine learning in solid-state materials science," *npj Computational Materials*, vol. 5, no. 1. Nature Publishing Group, pp. 1–36, Dec. 01, 2019, doi: 10.1038/s41524-019-0221-0.

[9] M. N. Gjerding *et al.*, "Recent progress of the computational 2D materials database (C2DB)," *2D Mater.*, vol. 8, no. 4, p. 044002, Jul. 2021, doi: 10.1088/2053-1583/AC1059.

[10] S. Masubuchi *et al.*, "Deep-learning-based image segmentation integrated with optical microscopy for automatically searching for two-dimensional materials," *npj 2D Mater. Appl.*, vol. 4, no. 1, pp. 1–9, Dec. 2020, doi: 10.1038/s41699-020-0137-z.

[11] Y. Liu, T. Zhao, W. Ju, S. Shi, S. Shi, and S. Shi, "Materials discovery and design using machine learning," *J. Mater.*, vol. 3, no. 3, pp. 159–177, Sep. 2017, doi: 10.1016/J.JMAT.2017.08.002.

[12] Y. Li *et al.*, "Rapid identification of two-dimensional materials via machine learning assisted optic microscopy," *J. Mater.*, vol. 5, no. 3, pp. 413–421, Sep. 2019, doi: 10.1016/j.jmat.2019.03.003.

[13] T. K. Patra *et al.*, "Defect dynamics in 2-D MoS2 probed by using machine learning, atomistic simulations, and high-resolution microscopy," *ACS Nano*, vol. 12, no. 8, pp. 8006–8016, Aug.

2018, doi: 10.1021/acsnano.8b02844.

[14] X. Lin *et al.*, “Intelligent identification of two-dimensional nanostructures by machine-learning optical microscopy,” *Nano Res.*, vol. 11, no. 12, pp. 6316–6324, Dec. 2018, doi: 10.1007/s12274-018-2155-0.

[15] K. Ghosal and K. Sarkar, “Biomedical Applications of Graphene Nanomaterials and beyond,” *ACS Biomaterials Science and Engineering*, vol. 4, no. 8. American Chemical Society, pp. 2653–2703, Aug. 13, 2018, doi: 10.1021/acsbiomaterials.8b00376.

[16] B. N. Shivananju, H. Y. Hoh, W. Yu, and Q. Bao, “Optical biochemical sensors based on 2D materials,” in *Fundamentals and Sensing Applications of 2D Materials*, Elsevier, 2019, pp. 379–406.

[17] J. Yi, G. Choe, J. Park, and J. Y. Lee, “Graphene oxide-incorporated hydrogels for biomedical applications,” *Polymer Journal*, vol. 52, no. 8. Springer Nature, pp. 823–837, Aug. 01, 2020, doi: 10.1038/s41428-020-0350-9.

[18] G. Chilkoor *et al.*, “Hexagonal Boron Nitride for Sulfur Corrosion Inhibition,” *ACS Nano*, vol. 14, no. 11, pp. 14809–14819, Nov. 2020, doi: 10.1021/ACSNANO.0C03625.

[19] G. Chilkoor *et al.*, “Maleic anhydride-functionalized graphene nanofillers render epoxy coatings highly resistant to corrosion and microbial attack,” *Carbon N. Y.*, vol. 159, pp. 586–597, Apr. 2020, doi: 10.1016/J.CARBON.2019.12.059.

[20] S. Kumar and K. Chatterjee, “Comprehensive Review on the Use of Graphene-Based Substrates for Regenerative Medicine and Biomedical Devices,” *ACS Applied Materials and Interfaces*, vol. 8, no. 40. American Chemical Society, pp. 26431–26457, Oct. 12, 2016, doi: 10.1021/acsami.6b09801.

[21] M. C. Demirel, M. Vural, and M. Terrones, “Composites of Proteins and 2D Nanomaterials,” *Adv. Funct. Mater.*, vol. 28, no. 27, p. 1704990, Jul. 2018, doi: 10.1002/ADFM.201704990.

[22] S. Kumar and K. Chatterjee, “Comprehensive Review on the Use of Graphene-Based Substrates for Regenerative Medicine and Biomedical Devices,” *ACS Appl. Mater. Interfaces*, vol. 8, no. 40, pp. 26431–26457, Oct. 2016, doi: 10.1021/ACSAMI.6B09801.

[23] M. G. Rasul, A. Kiziltas, B. Arfaei, and R. Shahbazian-Yassar, “2D boron nitride nanosheets for polymer composite materials,” *npj 2D Mater. Appl. 2021 51*, vol. 5, no. 1, pp. 1–18, May 2021, doi: 10.1038/s41699-021-00231-2.

[24] S. Li, K. Yang, C. Tan, X. Huang, W. Huang, and H. Zhang, “Preparation and applications of novel composites composed of metal–organic frameworks and two-dimensional materials,” *Chem. Commun.*, vol. 52, no. 8, pp. 1555–1562, Jan. 2016, doi: 10.1039/C5CC09127F.

[25] I. Demeridou *et al.*, “2D semiconducting materials for electronic and optoelectronic applications: potential and challenge Recent citations,” *2D Mater.*, 2018, doi: 10.1088/2053-1583/ab6267.

[26] F. Withers *et al.*, “Light-emitting diodes by band-structure engineering in van der Waals heterostructures,” *Nat. Mater.*, vol. 14, no. 3, pp. 301–306, Feb. 2015, doi: 10.1038/nmat4205.

[27] M. F. El-Kady, Y. Shao, and R. B. Kaner, “Graphene for batteries, supercapacitors and beyond,” *Nature Reviews Materials*, vol. 1, no. 7. Nature Publishing Group, pp. 1–14, May 24, 2016, doi:

10.1038/natrevmats.2016.33.

[28] J. Islam, G. Chilkoor, K. Jawaharraj, S. S. Dhiman, R. Sani, and V. Gadhamshetty, "Vitamin-C-enabled reduced graphene oxide chemistry for tuning biofilm phenotypes of methylotrophs on nickel electrodes in microbial fuel cells," *Bioresour. Technol.*, vol. 300, p. 122642, Mar. 2020, doi: 10.1016/J.BIORTECH.2019.122642.

[29] S. Homaeigohar and M. Elbahri, "Graphene membranes for water desalination," *NPG Asia Mater.*, vol. 9, no. 8, pp. e427–e427, Aug. 2017, doi: 10.1038/am.2017.135.

[30] O. Moldovan, B. Iñiguez, M. J. Deen, and L. F. Marsal, "Graphene electronic sensors - Review of recent developments and future challenges," *IET Circuits, Devices and Systems*, vol. 9, no. 6. Institution of Engineering and Technology, pp. 446–453, Nov. 01, 2015, doi: 10.1049/iet-cds.2015.0259.

[31] P. P. Mane, R. S. Ambekar, and B. Kandasubramanian, "Electrospun nanofiber-based cancer sensors: A review," *International Journal of Pharmaceutics*, vol. 583. Elsevier B.V., Jun. 15, 2020, doi: 10.1016/j.ijpharm.2020.119364.

[32] L. Zhou *et al.*, "Nonlinear optical characterization of 2d materials," *Nanomaterials*, vol. 10, no. 11. MDPI AG, pp. 1–38, Nov. 01, 2020, doi: 10.3390/nano10112263.

[33] J. Dan, X. Zhao, and S. J. Pennycook, "A machine perspective of atomic defects in scanning transmission electron microscopy," *InfoMat*, vol. 1, no. 3, pp. 359–375, Sep. 2019, doi: 10.1002/inf2.12026.

[34] D. D. M. Agany, J. E. Pietri, and E. Z. Gnimpieba, "Assessment of vector-host-pathogen relationships using data mining and machine learning," *Comput. Struct. Biotechnol. J.*, vol. 18, pp. 1704–1721, Jan. 2020, doi: 10.1016/J.CSBJ.2020.06.031.

[35] R. M. Sterbentz, K. L. Haley, and J. O. Island, "Universal image segmentation for optical identification of 2D materials," *Sci. Rep.*, vol. 11, no. 1, Dec. 2021, doi: 10.1038/s41598-021-85159-9.

[36] M. Y. & K. T. Yu Saito, Kento Shin, Kei Terayama, Shaan Desai, Masaru Onga, Yuji Nakagawa, Yuki M. Itahashi, Yoshihiro Iwasa, "Deep-learning-based quality filtering of mechanically exfoliated 2D crystals," *npj Comput. Mater.*, 2019, [Online]. Available: https://www.nature.com/articles/s41524-019-0262-4.

[37] S. Masubuchi and T. Machida, "Classifying optical microscope images of exfoliated graphene flakes by data-driven machine learning," *npj 2D Mater. Appl.*, vol. 3, no. 1, Dec. 2019, doi: 10.1038/s41699-018-0084-0.

[38] B. Han *et al.*, "Deep-Learning-Enabled Fast Optical Identification and Characterization of 2D Materials," *Adv. Mater.*, vol. 32, no. 29, Jun. 2020, doi: 10.1002/adma.202000953.

[39] J. Yang and H. Yao, "Automated identification and characterization of two-dimensional materials via machine learning-based processing of optical microscope images," *Extrem. Mech. Lett.*, vol. 39, Sep. 2020, doi: 10.1016/j.eml.2020.100771.

[40] M. J. Abedin, T. Barua, M. Shaibani, and M. Majumder, "A High Throughput and Unbiased Machine Learning Approach for Classification of Graphene Dispersions," *Adv. Sci.*, vol. 7, no. 20, pp. 1–11, Oct. 2020, doi: 10.1002/advs.202001600.

[41] J. A. Aguiar, M. L. Gong, R. R. Unocic, T. Tasdizen, and B. D. Miller, "Decoding crystallography from high-resolution electron imaging and diffraction datasets with deep learning," *Sci. Adv.*, vol. 5, no. 10, Oct. 2019, doi: 10.1126/sciadv.aaw1949.

[42] R. Mizuguchi, Y. Igarashi, H. Imai, and Y. Oaki, "Lateral-size control of exfoliated transition-metal-oxide 2D materials by machine learning on small data," *Nanoscale*, vol. 13, no. 6, pp. 3853–3859, Feb. 2021, doi: 10.1039/d0nr08684c.

[43] K. Noda, Y. Igarashi, H. Imai, and Y. Oaki, "Efficient Syntheses of 2D Materials from Soft Layered Composites Guided by Yield Prediction Model: Potential of Experiment-Oriented Materials Informatics," *Adv. Theory Simulations*, vol. 3, no. 7, Jul. 2020, doi: 10.1002/adts.202000084.

[44] T. Ahmed *et al.*, "Fully Light-Controlled Memory and Neuromorphic Computation in Layered Black Phosphorus," *Adv. Mater.*, vol. 33, no. 10, Mar. 2021, doi: 10.1002/adma.202004207.

[45] J. Chen *et al.*, "Computer vision AC-STEM automated image analysis for 2D nanopore applications," *Ultramicroscopy*, p. 113249, Mar. 2021, doi: 10.1016/j.ultramic.2021.113249.

[46] Z. Maxim, S. Jesse, B. G. Sumpter, S. V. Kalinin, and O. Dyck, "Tracking atomic structure evolution during directed electron beam induced Si-atom motion in graphene via deep machine learning," *Nanotechnology*, vol. 32, no. 3, Jan. 2021, doi: 10.1088/1361-6528/abb8a6.

[47] M. R. Rashidian Vaziri, "High-contrast optical microscopy of graphene sheets," *Microsc. Res. Tech.*, vol. 83, no. 9, pp. 1132–1140, Sep. 2020, doi: 10.1002/jemt.23505.

[48] M. P. V. Christiansen, H. L. Mortensen, S. A. Meldgaard, and B. Hammer, "Gaussian representation for image recognition and reinforcement learning of atomistic structure," *J. Chem. Phys.*, vol. 153, no. 4, Jul. 2020, doi: 10.1063/5.0015571.

[49] E. Ertekin *et al.*, "Crowd-sourced data and analysis tools for advancing the chemical vapor deposition of graphene: Implications for manufacturing," *ACS Appl. Nano Mater.*, vol. 3, no. 10, pp. 10144–10155, Oct. 2020, doi: 10.1021/acsanm.0c02018.

[50] L. Mennel, J. Symonowicz, S. Wachter, D. K. Polyushkin, A. J. Molina-Mendoza, and T. Mueller, "Ultrafast machine vision with 2D material neural network image sensors," *Nature*, vol. 579, no. 7797, pp. 62–66, Mar. 2020, doi: 10.1038/s41586-020-2038-x.

[51] F. Cellini, F. Lavini, C. Berger, W. De Heer, and E. Riedo, "Layer dependence of graphene-diamene phase transition in epitaxial and exfoliated few-layer graphene using machine learning," *2D Mater.*, vol. 6, no. 3, Jun. 2019, doi: 10.1088/2053-1583/ab1b9f.

[52] M. Ziatdinov, O. Dyck, S. Jesse, and S. V. Kalinin, "Atomic Mechanisms for the Si Atom Dynamics in Graphene: Chemical Transformations at the Edge and in the Bulk," *Adv. Funct. Mater.*, vol. 29, no. 52, Dec. 2019, doi: 10.1002/adfm.201904480.

[53] M. Ziatdinov *et al.*, "Deep Learning of Atomically Resolved Scanning Transmission Electron Microscopy Images: Chemical Identification and Tracking Local Transformations," *ACS Nano*, vol. 11, no. 12, pp. 12742–12752, Dec. 2017, doi: 10.1021/acsnano.7b07504.

[54] N. Laanait, M. Ziatdinov, Q. He, and A. Borisevich, "Identifying local structural states in atomic imaging by computer vision," *Adv. Struct. Chem. Imaging*, vol. 2, no. 1, p. 14, Jan. 2016, doi: 10.1186/s40679-016-0028-8.

[55] Q. Luo, E. A. Holm, and C. Wang, “A transfer learning approach for improved classification of carbon nanomaterials from TEM images,” *Nanoscale Adv.*, vol. 3, no. 1, pp. 206–213, Jan. 2021, doi: 10.1039/d0na00634c.

[56] P. Hundi and R. Shahsavari, “Deep Learning to Speed up the Development of Structure–Property Relations For Hexagonal Boron Nitride and Graphene,” *Small*, vol. 15, no. 19, May 2019, doi: 10.1002/smll.201900656.

[57] A. Trentino *et al.*, “Atomic-Level Structural Engineering of Graphene on a Mesoscopic Scale.,” *Nano Lett.*, p. acs.nanolett.1c01214, Jun. 2021, doi: 10.1021/acs.nanolett.1c01214.

[58] Y. J. Shin, W. Shin, T. Taniguchi, K. Watanabe, P. Kim, and S. H. Bae, “Fast and accurate robotic optical detection of exfoliated graphene and hexagonal boron nitride by deep neural networks,” *2D Mater.*, vol. 8, no. 3, Jul. 2021, doi: 10.1088/2053-1583/abd72c.

[59] K.-C. Li *et al.*, “Intelligent Identification of MoS_2 Nanostructures with Hyperspectral Imaging by 3D-CNN,” *Nanomater. 2020, Vol. 10, Page 1161*, vol. 10, no. 6, p. 1161, Jun. 2020, doi: 10.3390/NANO10061161.

[60] K. Momeni *et al.*, “Multiscale computational understanding and growth of 2D materials: a review,” *npj Comput. Mater. 2020 61*, vol. 6, no. 1, pp. 1–18, Mar. 2020, doi: 10.1038/s41524-020-0280-2.