

AUTOMATIC 2D MATERIAL DETECTION AND QUANTUM EMISSION
PREDICTION USING DEEP LEARNING-BASED MODELS

by

Fereshteh Ramezani

A dissertation submitted in partial fulfillment
of the requirements for the degree

of

Doctor of Philosophy

in

Electrical Engineering

MONTANA STATE UNIVERSITY
Bozeman, Montana

December 2023

©COPYRIGHT

by

Fereshteh Ramezani

Creative Commons Attribution 4.0 International License 

DEDICATION

I dedicate this work to my family: my mother and father, my grandparents, and my sisters who have always been a source of motivation and support for me. This journey would not have been possible without them.

ACKNOWLEDGEMENTS

I extend my heartfelt gratitude to my dedicated advisor, Dr. Bradley Whitaker, for his unwavering support and guidance from the very beginning of my Ph.D. journey. I also want to convey my deepest appreciation to Dr. Nicholas Borys, the Associate Director of MonArk Quantum Foundry, for his support at every stage of my research. Additionally, I would like to express my sincere gratitude to the esteemed members of my committee, Dr. Robert Maher and Dr. Brock LaMeres, for their valuable feedback and contributions to my work. I acknowledge the MonArk NSF Quantum Foundry supported by the National Science Foundation Q-AMASE-i program under NSF award No. DMR-1906383.

TABLE OF CONTENTS

1. INTRODUCTION	1
Quantum Materials and Properties	1
Machine Learning for Detection and Prediction	2
Dissertation Outline.....	3
2. BACKGROUND.....	6
2D Material.....	6
Atomic Structure and Application	6
Material Exfoliation.....	8
Quantum Fluctuations.....	11
Object Detection	12
Traditional Methods	14
Supervised Machine Learning	18
Unsupervised Learning	24
Applications of Object Detection Methods	26
Time-series Forecasting	27
Recurrent Neural Networks	28
Correlation and Autocorrelation.....	31
3. 2D MATERIAL DETECTION	33
ML for 2D Material Detection	33
Traditional Image Processing Methods	33
Supervised Learning Methods.....	37
Unsupervised Learning Methods.....	46
Methodology	47
Data Acquisition	47
Data Labeling	48
Model and Training.....	49
Results.....	51
Discussion	59
Integrated 2D Flake Detection.....	60
4. QUANTUM EMISSION PREDICTION.....	66
Fluctuations of QEs in 1L-WS ₂	68
Methodology	69
Data Acquisition	69
Preprocessing.....	70

TABLE OF CONTENTS – CONTINUED

Experiment and Model	71
Evaluation	72
Results.....	72
Correlation	72
Prediction.....	73
Discussion	83
5. CONCLUSION	85
REFERENCES CITED.....	87

LIST OF TABLES

Table	Page
3.1 Dataset Labels and Annotations	49
4.1 Root Mean Square Error using LSTM and Shifted Forecast	76
4.2 Root Mean Square Error Percentage using LSTM and Shifted Forecast	77

LIST OF FIGURES

Figure	Page
2.1 Different 2D Materials [1].	7
2.2 Mechanical exfoliation of 2D crystals. (a) Pressing adhesive tape against 2D crystal. (b) Top few layers are attached to the tape. (c) Pressing the tape with crystals against a surface of choice. (d) The bottom layer remains on the substrate after peeling off [2].	9
2.3 Optical microscopic images after exfoliation. (a) hBN, (b) MoS ₂ , (c) Graphene and (d) WTe ₂ [3].	10
2.4 Fluctuations in intensity and energy for a nanobubble QE in 1L-WS ₂ on an Au surface.	12
2.5 Architecture of R-CNN object detection [4].	20
2.6 Fast R-CNN framework for object detection [5].	21
2.7 Architecture of Mask R-CNN framework [6].	22
2.8 YOLO framework [7].	23
2.9 RetinaNet network architecture [8].	24
2.10 Recurrent neural network and feed-forward neural network architectures. The figure on the left shows a recurrent neural network with recurrent connections on hidden states which do the back propagation, whereas the figure on the right shoes the feed-forward architecture of neural networks that have no recurrent connections.	29
2.11 Unfolded sequential architecture in an RNN. X_t is the input at time step t, Y_t is the output at time step t and H_t is the hidden state at time t.	30
2.12 Memory cells structure in LSTM [9].	31

LIST OF FIGURES – CONTINUED

Figure	Page
3.1 Proposed deep learning pipeline for 2D fake detection. First, a dataset of hBN samples are collected through mechanical exfoliation. Then as the second step, the dataset is annotated manually using an online annotation tool. During the third step, DetectoRS is chosen as the training algorithm, and for the last step, the trained model is applied to the test images to detect hBN samples.	47
3.2 Architecture of Switchable Atrous Convolution (SAC) [10].....	50
3.3 Results obtained from using the proposed algorithm on one image from our dataset; Part (a) shows the raw image, part (b) shows the labels generated by the algorithm, and parts (c) and (d) are zoomed-in regions showing the flake type and confidence score reported by the algorithm. We manually identified 19 flakes in the image, and the algorithm correctly classified 14 of them.	54
3.4 Confusion matrix template for a generic multi-class classification problem, showing how precision and recall are calculated for each class.....	55
3.5 Confusion matrix showing results obtained by applying the proposed algorithm to all three test datasets (MSU, Masubuchi, and UA).	56
3.6 Results obtained from using the proposed algorithm on one image from the Masubuchi dataset; Part (a) shows the raw image, part (b) shows the labels generated by the algorithm, and parts (c) and (d) are zoomed-in regions showing the flake type and confidence score reported by the algorithm. We manually identified 16 flakes in the image, and the algorithm correctly classified 12 of them (microscope images obtained from https://doi.org/10.6084/m9.figshare.11881053 under the license CC BY 4.0)	57

LIST OF FIGURES – CONTINUED

Figure	Page
3.7 Results obtained from using the proposed algorithm on one image from the UA dataset; Part (a) shows the raw image, part (b) shows the labels generated by the algorithm, and parts (c) and (d) are zoomed-in regions showing the flake type and confidence score reported by the algorithm. We manually identified 32 flakes in the image, and the algorithm correctly classified 20 of them. (Microscope images obtained by Hugh Churchill and used with permission.).....	58
3.8 Confusion matrices showing results when applying Masubuchi's pre-trained Mask R-CNN to our all three test datasets (MSU, Masubuchi, UA). The pre-trained model was unable to identify most manually labeled flakes in the test images.	59
3.9 Makesense.ai is a free-to-use online tool for labeling photos.....	62
3.10 VGG Image Annotator is a simple and standalone manual annotation software for image, audio and video.....	63
3.11 Real-time 2D material detection.....	65
4.1 Fluctuations in intensity and energy for a nanobubble QE in single-layer WS ₂ on an Au surface. The figure on the left is shows the number of photons emitted by the device as a function of time and photon wavelength. The figure on the right shows the same data, but also indicates four different wavelength bands where quantum emissions occur.....	69
4.2 Correlation matrix and autocorrelation plot for band 1 and its center for 30 lags. Image on the left represents the correlation matrix for wavelengths withing band 1 and the image on the right represents the autocorrelation lambda = 606.49 nm.	73
4.3 Correlation matrix and autocorrelation plot for band 2 and its center for 30 lags. Image on the left represents the correlation matrix for wavelengths withing band 1 and the image on the right represents the autocorrelation lambda=613.77 nm.	74

LIST OF FIGURES – CONTINUED

Figure	Page
4.4 Correlation matrix and autocorrelation plot for band 3 and its center for 30 lags. Image on the left represents the correlation matrix for wavelengths withing band 1 and the image on the right represents the autocorrelation lambda=620.38 nm.	75
4.5 Correlation matrix and autocorrelation plot for band 4 and its center for 30 lags. Image on the left represents the correlation matrix for wavelengths withing band 1 and the image on the right represents the autocorrelation lambda=630.37 nm.	76
4.6 Prediction results using LSTM and shifted forecasts for the full image where all the wavelengths within the dataset are considered. Columns represent the actual image, forecast method, and difference between the actual image and the forecast. The first row represents the LSTM model and the second row represents the shifted forecast.	78
4.7 Prediction results using LSTM and shifted forecasts for band 1. Columns represent the actual image, forecast method, and difference between the actual image and the forecast. The first row represents the LSTM model and the second row represents the shifted forecast.	79
4.8 Prediction results using LSTM and shifted forecasts for band 2. Columns represent the actual image, forecast method, and difference between the actual image and the forecast. The first row represents the LSTM model and the second row represents the shifted forecast.	80
4.9 Prediction results using LSTM and shifted forecasts for band 3. Columns represent the actual image, forecast method, and difference between the actual image and the forecast. The first row represents the LSTM model and the second row represents the shifted forecast.	81

LIST OF FIGURES – CONTINUED

Figure	Page
4.10 Prediction results using LSTM and shifted forecasts for band 4. Columns represent the actual image, forecast method, and difference between the actual image and the forecast. The first row represents the LSTM model and the second row represents the shifted forecast.	82
4.11 Single step prediction for wavelength 620.52 nm and 619.17 nm using LSTM forecast and Shifted forecast. For 620.52 nm, LSTM and shifted forecast RMSE are 420.37 and 346.7. For 619.17 nm, LSTM and shifted forecast RMSE are 648.80 and 654.57 respectively.	83

ABSTRACT

The realm of quantum engineering holds immense promise for revolutionizing technological landscapes, particularly with the advent of 2D materials in quantum device applications. The fundamental properties of these materials make them pivotal in various quantum applications. However, the progress in quantum engineering faces significant roadblocks, primarily centered around two challenges: accurate 2D material detection and understanding the random nature of quantum fluctuations. In response to the first challenge, I have successfully implemented a new deep learning pipeline to identify 2D materials in microscopic images. I have used a state-of-the-art two-stage object detector and trained it on images containing flakes of varying thickness of hexagonal boron nitride (hBN, a 2D material). The trained model achieved a high detection accuracy for the rare category of thin flakes (≤ 50 atomic layers thick). My further analysis shows that this proposed pipeline is robust against changes in color or substrate background, and could be generalized to various microscope settings. As an achievement, I have integrated my proposed method to the 2D quantum material pipeline (2D-QMaP), that has been under development by the MonArk Quantum Foundry, to provide automated capabilities that unite and accelerate the primary stages of sample preparation and device fabrication for 2D quantum materials research. My proposed algorithm has given the 2D-QMaP fully automatic real-time 2D flake detection capabilities, which has never been done effectively before. To address the second challenge, I assessed the random nature of quantum fluctuations, and I developed time series forecasting deep learning models to analyze and predict quantum emission fluctuations for the first time. My trained models can roughly follow the actual trend of the data and, under certain data processing conditions, can predict peaks and dips of the fluctuations. The ability to anticipate these fluctuations will allow physicists to harness quantum fluctuation characteristics to develop novel scientific advances in quantum computing that will greatly benefit quantum technologies. The automated 2D material identification, addressing the laborious process of flake detection, and the introduction of innovative quantum fluctuations analysis with predictive capabilities not only streamline research processes but also hold the promise of creating more stable and dependable quantum emission devices, thus significantly advancing the broader field of quantum engineering.

INTRODUCTION

Quantum Materials and Properties

Two-dimensional (2D) materials, consisting of one or a few atomic layers in their physical structure, exhibit many quantum phenomena in systems. Exploring these features, including strong many-body interactions, pristine interfaces, and strong confinement in a single direction, may lead to many research and development opportunities [11–14]. In bulk, 2D materials form van der Waals crystals where strong in-plane bonding within a single layer is complemented by significantly weaker van der Waals bonding between the layers [15]. Van der Waals forces refer to the attractive or repulsive forces between molecules or atoms. These forces play a crucial role in the behavior of gases, liquids, and solids. Van der Waals crystals typically refer to crystals held together primarily by van der Waals forces. The anisotropy in bonding strengths, combined with a chemical structure where the bonds within a single layer are completely passivated, enables individual layers of the 2D material to be isolated, forming chemically and structurally robust sheets of atoms that are atomically thick (i.e., 1-3 atomic layers thick). Although ultrathin, these 2D crystallites can have lateral sizes that are $10\text{-}100\text{ }\mu\text{m} - 10,000\text{-}100,000\times$ larger than their thickness.

In single- and few-layer form, many 2D materials exhibit excellent optical, electronic, and/or magnetic phenomena that provide means to prepare, interact with, and study quantum states in matter [16–20]. With over 1000 known van der Waals 2D materials that span all functionalities (insulators, metals, semiconductors, ferromagnets, ferroelectrics, etc.), a prolific parameter space of systems and structure-property relationships for quantum, optoelectronic, and magnetic technologies exists [21–23]. Different Van der Waals materials have been used to demonstrate appealing device applications such as emitting diodes. One

such material is hexagonal boron nitride (hBN).

The wavelength, polarization, and intensity of the photons from a quantum emitter must be stable in order for integrated quantum photonics to be as useful and valuable as possible. However, random variation of emission energy caused by the inhomogeneity in local environment is a major challenge for all solid-state single photon emitters. The task of exploring the vast world of 2D materials requires the ability to efficiently and reliably produce single- and few-layer samples of 2D materials [24, 25]. However, the most widely used state-of-the-art method for preparing single- and few-layer samples of 2D materials relies on the manual exfoliation of a large population of smaller crystallites from a bulk crystal. The resulting crystallites have a range of thicknesses, and the entire population must be manually searched using optical reflected light microscopy to identify those that have a desired thickness, which is most often the thickness of a single layer. This labor-intensive method for collecting single- and few-layer samples strongly inhibits rapid discovery and exploration of these material systems.

Machine Learning for Detection and Prediction

Object detection is an important computer vision task that deals with detecting instances of visual objects of certain categories in digital images. The goal of object detection is to develop computational models and techniques that provide one of the most basic pieces of information needed by computer vision applications [26]. In other words, the goal of object detection is to determine whether there are any instances of objects from given categories in an image and, if present, to return the spatial location and extent of each object instance [27]. Object detection supports a wide range of applications, including robot vision, consumer markets, autonomous driving, human computer interaction, content based image retrieval, intelligent video surveillance, and augmented reality [27, 28].

Rapid progressions in deep learning (DL) and improvements in device capabilities

including computing power, memory capacity, power consumption, image sensor resolution, and optics have improved the performance and cost-effectiveness and assisted the spread of vision-based applications as well as time-series problems. Compared to traditional computer vision techniques, which have a long trial-and-error process, deep learning enables end-to-end object detectors to achieve greater accuracy in tasks such as image classification, semantic segmentation, object detection and Simultaneous Localization and Mapping [29]. Deep learning computer vision methods identify and analyze numerous images into useful information at a fast rate with a high accuracy. Deploying DL algorithms for 2D material search can automate the flake detection task with minimal need for human intervention. So far, there have been various attempts taken toward flake identification in images [3, 30–35]. However, these methods failed to detect flakes within different microscope settings and image magnification.

Time-series modeling has held a significant role in academic research and practical applications across various fields, including climate modeling [36], biological sciences[37], medicine[38], retail[39], and finance[40]. Traditional approaches have centered on parametric models informed by domain knowledge, such as autoregressive (AR), exponential smoothing, or structural time-series models. However, modern machine learning techniques offer the capability to capture temporal patterns solely through data-driven methods. Given the growing availability of data and improved computing power in recent years, deep learning has emerged as a crucial component of the next wave of time-series forecasting models. [41]. Therefore, deep learning time-series forecasting models have the capacity and potential to analyze and address QE prediction problem.

Dissertation Outline

In this work, I focus on two obstacles preventing the commercial adoption of quantum materials: 2D material detection and quantum emitter fluctuation prediction. Solving this

problems will enable the mass production of quantum materials as well as improve the consistency of their behavior within a fabricated quantum device. In Chapter 2, I delve into the foundational knowledge required for a comprehensive understanding of the topics covered in this study. This chapter gives an overview of several key concepts, each contributing to the overall context and groundwork for my research. In this chapter, the nuances of computer vision techniques and object detection methods are explored, showcasing their diverse applications across domains. Firstly, I delve into the atomic-level intricacies of 2D materials, exploring their practical applications, and the fascinating quantum phenomena that underpin their behavior. Then, I dive into object detection and the talk about the traditional techniques versus machine learning techniques that are used in this area. Here, I explain the crucial role of machine learning in object detection and image processing. Furthermore, I investigate the art of time-series forecasting, underpinned by recurrent neural networks, highlighting its relevance in predicting future trends and patterns. This chapter serves as the intellectual bridge that connects the areas of machine learning and quantum material development and characterization, setting the stage for the innovative research that follows.

In Chapter 3, I present on a comprehensive exploration of 2D materials detection. This chapter presents an overview of prior research in the field, providing insights into established methodologies that laid the foundation for my work. Subsequently, I discuss my contribution in developing a new detection pipeline which not only makes improvements in detection accuracy, but also is able to detect the rare class of thin flakes (< 50 atomic layers). My work is divided into several critical areas: data acquisition, data preprocessing and labeling, model development and training, results analysis, and the integration of techniques for a fully automatic real-time 2D flake detection. By examining the progress made in this domain and my innovative approaches, I contribute to the advancement of 2D material research and studies.

In continuing my research in machine learning for 2D materials, chapter 4 presents my work on developing time-series forecasting machine learning models to analyze and predict QE fluctuations for the first time. Within this chapter, I initiate my exploration by introducing the fundamental concepts and complexities surrounding this phenomenon. Here, I investigate the fluctuations of QEs in monolayer tungsten disulfide (1L-WS₂), and state the problem of QE fluctuations. Then I outline my methodology to address this problem, which includes data acquisition, preprocessing, and model development. Subsequently, I provide a detailed explanation of my experimental procedures and the models employed in predicting the QE patterns. As I navigate through this chapter, I offer comprehensive insights into my findings, encompassing correlations and the predictive capabilities of my models. Finally, I engage in a thorough discussion, highlighting the broader implications of my research within the intriguing field of quantum emission prediction.

Chapter 5 presents a comprehensive summary that distills the key insights and findings from the preceding chapters of this thesis. Over the course of the prior chapters, I have delved into a deep exploration of my extensive study on automatic detection of 2D materials and quantum emission prediction using deep learning based models, dissecting the various elements, uncovering critical insights, and examining the implications. This last chapter offers readers a consolidated perspective, allowing them to grasp the overarching themes and conclusions that have emerged from the earlier sections, and paving the way for future investigations in the field.

BACKGROUND

2D Material

Atomic Structure and Application

To study quantum phenomena in systems, researchers often turn to two-dimensional (2D) materials: single sheets of atoms arranged in a lattice. They are described as 2D because the sheets are much smaller in one dimension than the other two. Since they are only a few atomic layers thick and cannot be made thinner, 2D materials are sometimes called monolayers and are subject to the laws of quantum mechanics. Properties such as thermal and electrical conductivity are retained in a monolayer, but at this scale, the properties may become more or less efficient, or the material could take on other qualities altogether. These materials come in all types of different flavors. There are over 1,000 known 2D materials, each of which could possess a unique set of quantum properties. They can have unique band structures and characteristics which make them special for different applications.

Fig. 2.1 shows different 2D materials and their interesting applications areas. For example, graphene has distinctive band structure and high carrier mobilities for some applications, however it is not desirable for semiconductors due to its zero bandgap [1, 42]. Another 2D material, is hexagonal boron nitride (hBN), an isomorph of graphene with a very indistinguishable layered structure. hBN holds interesting opto-electrical properties merged with mechanical robustness, thermal stability and chemical inertness [43]. These properties have made hBN highly convenient for various demanding applications in Deep Ultraviolet (DUV) photonic devices [44], dielectric tunneling [45], power devices [46], electronic packaging [47], bio-medicines [48], and nanoelectronics [49]. Several 2D layered inorganic materials such as MoS_2 , WS_2 and SnS_2 , due to their layer-dependent electrical properties, have recently been developed and used in various applications like new sensors including gas sensing, electrochemical sensing, surface-enhanced Raman scattering (SERS)

and biosensing, SERS sensing and photodetection [50].

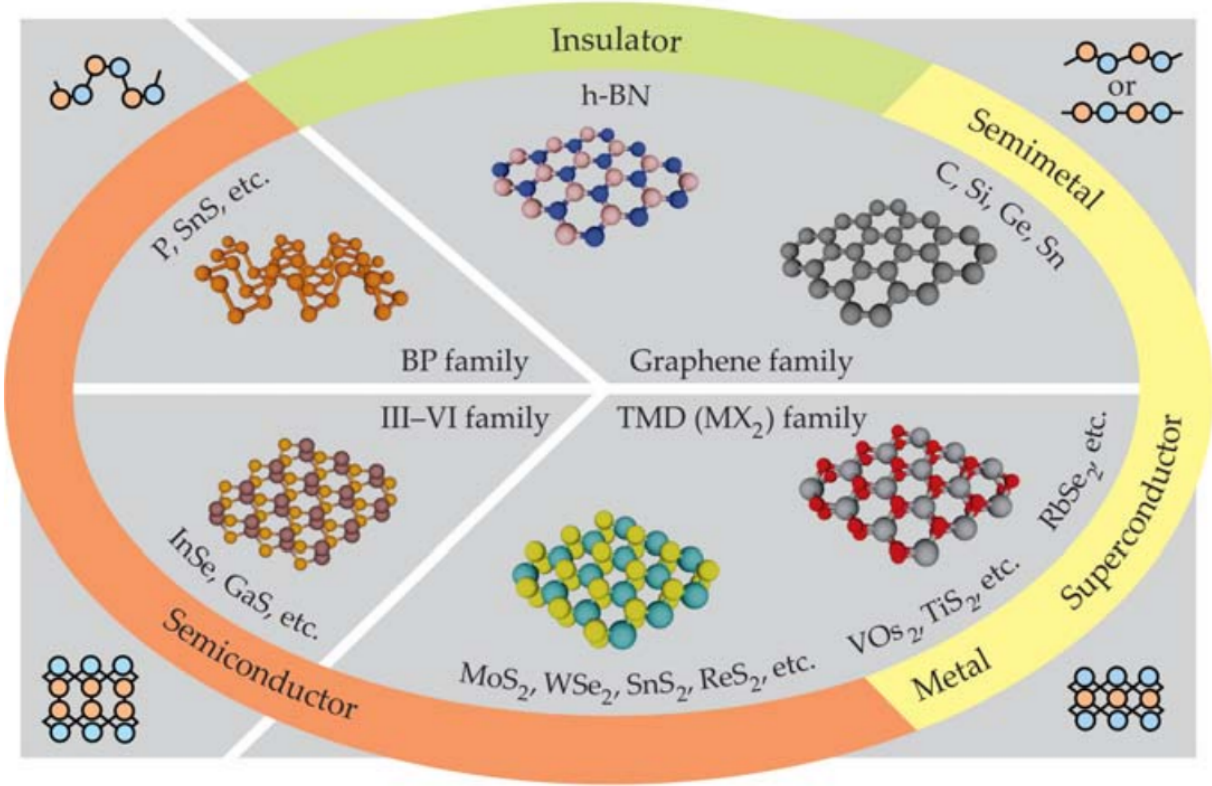


Figure 2.1: Different 2D Materials [1].

These 2D materials have been under research in diverse areas during the past several years [51–57]. Ultrathin 2D nanosheets of layered transition metal dichalcogenides (TMDs), unlike the graphene sheet, are chemically versatile, therefore they are of interest. Single- or few-layer TMDs are direct-gap semiconductors. Depending on their composition, structure and dimensionality, they can have different bandgap energies and carrier types (n- or p-type). In [58], Manish et al. have investigated the TMDs features which make them a good fit for various applications including opto-electronics and energy storage like batteries. Xu et al. [59] have investigated 2D materials considering their distinct properties and high speed surface areas for flexible supercapacitors as they are one of the most promising energy

storage devices. Wenbin et al. [60] have examined the thermodynamic and kinetic features of 2D phase transitions resulting from dimensionality confinement, elasticity, electrostatics, defects and chemistry inherent in 2D materials. They also discussed the applications of 2D phase transitions and future opportunities.

Material Exfoliation

The properties of 2D crystals discussed previously are very different from the 3D compounds from which they are derived. As an example, the 2D-material graphene is a zero-gap semiconductor, while the 3D-material compound graphite is a semimetal with a band overlap. Many monolayers of transition metal dichalcogenides (TMDC) in 2H phase are direct-bandgap semiconductors, whereas the band-gap of bulk forms is indirect. Any one of these materials, or layered combinations of them, have led to the creation of unique heterostructures, enabled the study of new physical effects and development of new devices [61] and also could be the basis of the next generation of quantum technology.

Crystalline planar structures bound by in-plane covalent bonds and out-of-plane van der Waals forces characterize layered materials. Single layers can be peeled away from bulk material by breaking the van der Waals bonds using adhesive tape [1]. The most effective methods to obtain high-quality single- or few-layer nanocrystals from their native multilayers structures are mechanical exfoliation and transfer [62]. Fig. 2.2 shows the mechanical exfoliation of 2D crystals using a tape, and Fig. 2.3 shows optical microscopic images from exfoliated 2D materials.

Relatively few of the 1,000-plus 2D materials have been explored because studying 2D materials has one major impediment: acquiring them in a single-layer form involves a time-consuming procedure. Hence, obtaining high quality samples of 2D materials is a major bottleneck. The manual exfoliation process based on transferring thin layers of material with adhesive tape is an impediment for future discovery and exploration of 2D materials [63], as

searching for these materials within microscopic images manually is a time consuming and tedious task.

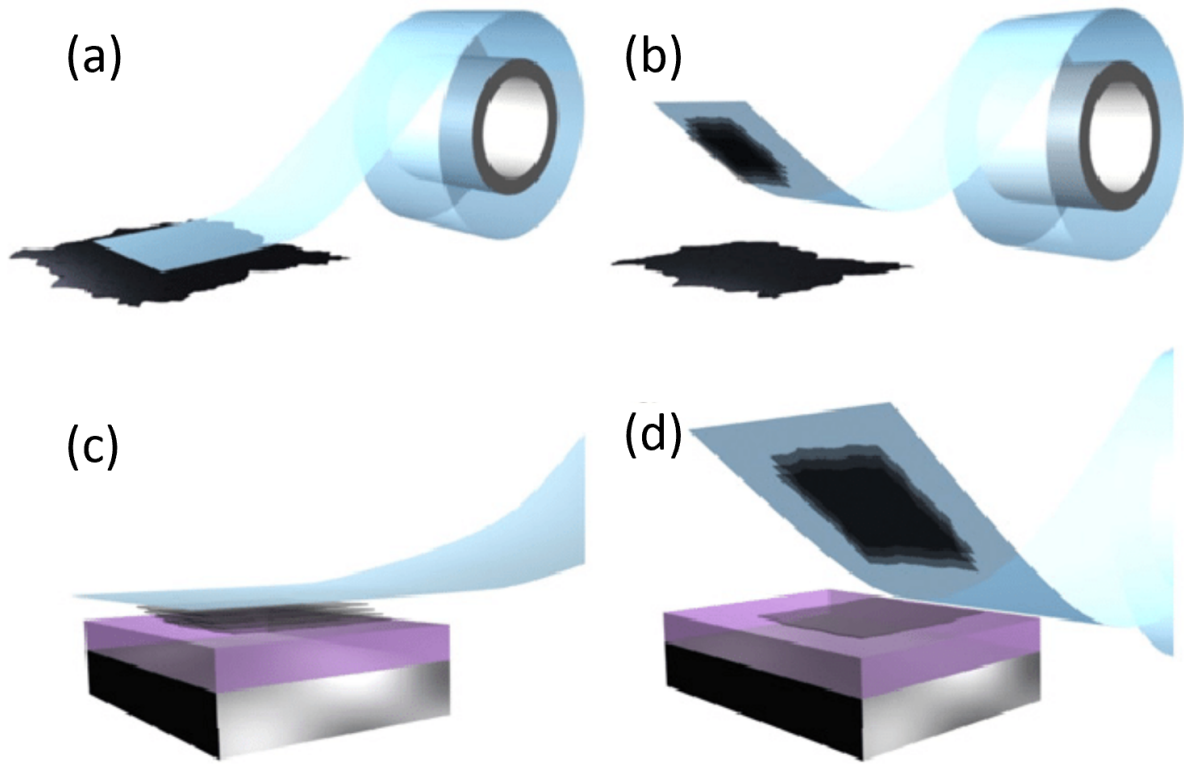


Figure 2.2: Mechanical exfoliation of 2D crystals. (a) Pressing adhesive tape against 2D crystal. (b) Top few layers are attached to the tape. (c) Pressing the tape with crystals against a surface of choice. (d) The bottom layer remains on the substrate after peeling off [2].

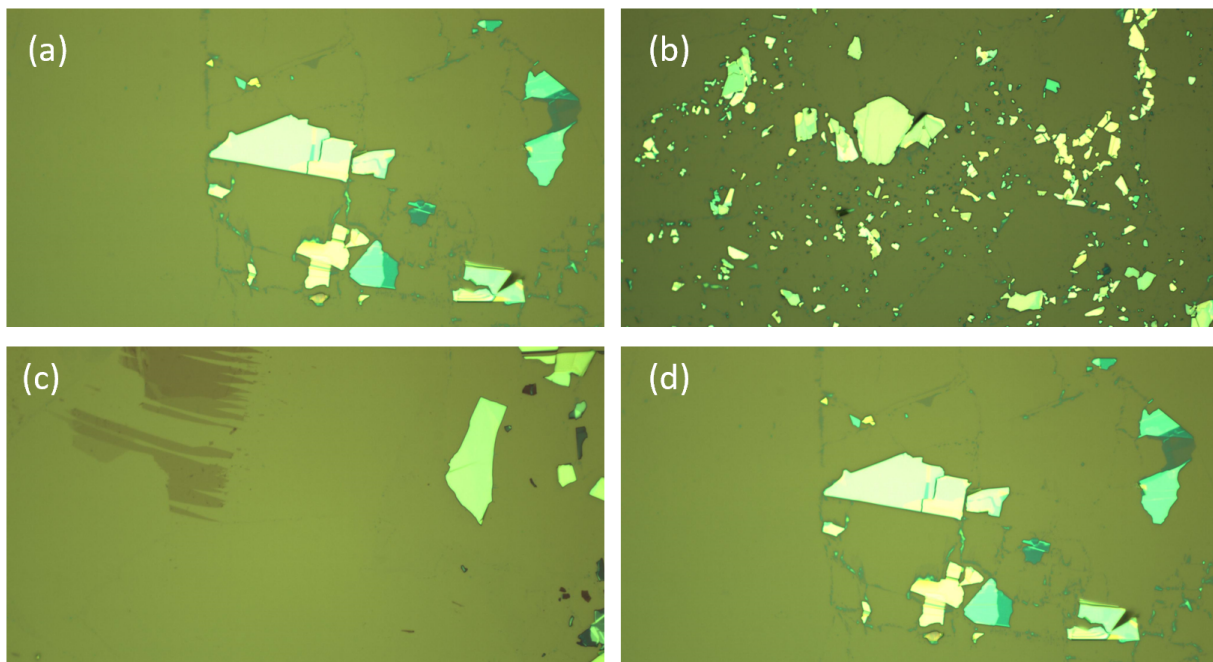


Figure 2.3: Optical microscopic images after exfoliation. (a) hBN, (b) MoS2, (c) Graphene and (d) WTe2 [3].

Quantum Fluctuations

Being able to apply quantum science to real world technologies would enormously benefit certain tasks such as communication [64], computation [65], simulation [66], sensing and fundamental science [67]. However, it requires novel manufacture of quantum technologies to be capable of implementing huge instances of quantum algorithms or deploying a global communication network. Exploiting single particles of light as quantum information carriers, together with wafer-scale fabrication, Integrated Quantum Photonics (IQP) is a compelling platform for the future of quantum technologies [68]. Various Van der Waals materials, including semi-metallic graphene, semiconducting transition metal dichalcogenides and insulating hBN, have been used to demonstrate attractive device applications [69] such as emitting diodes [70], laser [71] and optical modulators [72]. Quantum emitters (QEs) are the essential components in any integrated quantum photonic systems. To maximize the utility and value of IQP, the wavelength, polarization and intensity of the photons from a QE must be stable. However, the random variation of emission energy, caused by the inhomogeneity in the local environment, is a major challenge for all solid-state single photon emitters as it breaks the indistinguishability of single photons from multiple emitters [73].

Fig. 2.4 shows induced deleterious fluctuations in a QE caused by the local environment surrounding the QE, including nearby metallic and dielectric structures. Large-area encapsulation of the devices with thin layers of hBN provides a potential means to mitigate these fluctuations. However, they are not an absolute solution to eliminate them completely. To address further stability, QEs that are dynamically tunable can be used alongside a feedback system that intermittently samples of QLED emission and applies corrective stimuli. Machine learning time-series forecasting models have the potential to predict QE fluctuations and apply corrective stimuli to stabilize their emission wavelength and intensity.

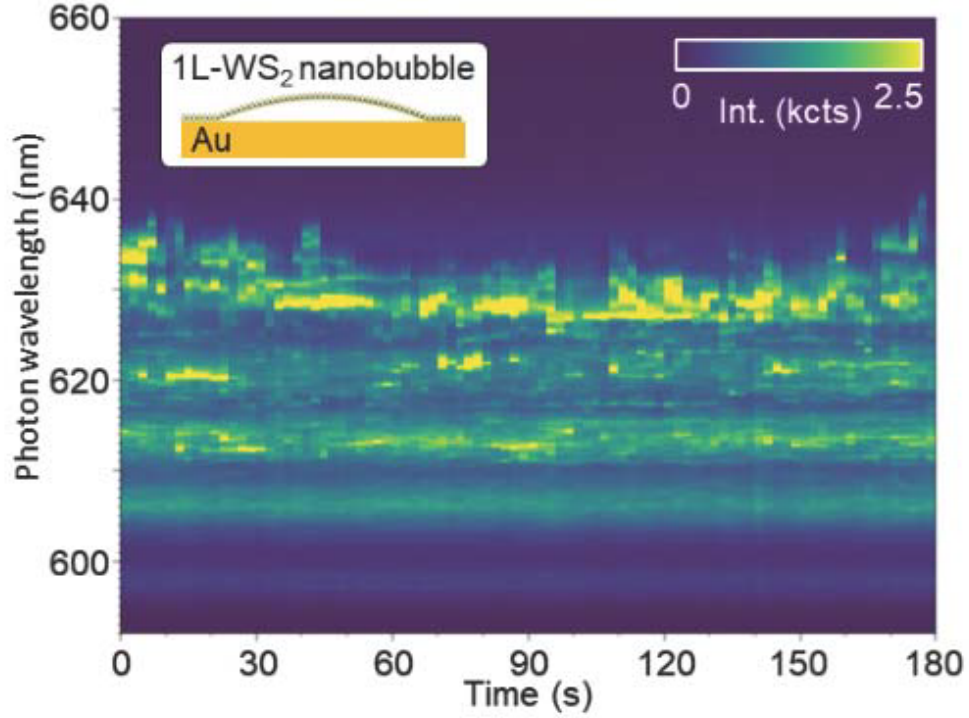


Figure 2.4: Fluctuations in intensity and energy for a nanobubble QE in 1L-WS₂ on an Au surface.

Object Detection

Computer vision is an interdisciplinary field that aims to enable computer-based systems to work the same as human vision, that is, to extract and understand useful information from visual inputs such as images and videos. There are various tasks in computer vision like object classification, segmentation and detection. Among current computer vision tasks, object detection is an important task that deals with detecting instances of visual objects of certain categories in digital images. The goal of object detection is to develop computational models and techniques that provide one of the most basic pieces of information needed by computer vision applications [26]. In other words, the goal of object detection is to determine

whether there are any instances of objects from given categories in an image and, if present, to return the spatial location and extent of each object instance [27]. Object detection supports a wide range of applications, including robot vision, consumer markets, autonomous driving, human computer interaction, content based image retrieval, intelligent video surveillance, and augmented reality [27, 28].

Over the past decades, object detection progress has mostly followed two historical eras: the traditional object detection period and the machine learning period. Traditional object detection methods deploy image processing techniques that require a domain expert to identify the majority of the applicable characteristics and features in order to simplify the data and make patterns easier for algorithms to learn. As a result, feature engineering is a major part of traditional methods where the raw data is needed to be selected, manipulated and transformed into features that can be used by the algorithms.

On the other hand, machine learning-based models have revolutionized object detection by automatically learning discriminative features from data. Machine learning is a subset of artificial intelligence that teaches computers to learn from experience. Without using a preexisting equation as a model, machine learning algorithms employ computer techniques to learn information directly from data. Machine learning-based approaches have grown dramatically over the past ten years, having an impact on a wide range of applications, including autonomous driving [74], healthcare [75], security systems [76], medicine [77], banking [78], and manufacturing [79], to name a few. Machine learning algorithms are capable of handling multivariate and multi-dimensional data, and of creating a mapping from the input data to decision-making outputs. They are able to classify the data into various categories without human intervention needed.

ML methods are commonly divided into four categories: supervised learning, unsupervised learning, semi-supervised learning and reinforcement learning. Supervised learning models have transformed object detection by leveraging labeled training data. These

models learn from labeled examples and require annotated data for training. In contrast, unsupervised learning-based models, like clustering or generative adversarial networks (GANs), can discover patterns and structures in unlabeled data. They aim to identify objects without explicit labels by learning from inherent data properties. Semi-supervised learning is situated between supervised and unsupervised learning, in that it uses both labeled and unlabeled data to perform specific learning tasks. The idea of these techniques is to leverage the large amount of unlabeled data available for many use cases, combined with a smaller set of labeled data [80]. In reinforcement learning, an agent receives a state and a reward from the environment. The agent’s job is to choose the best course of action that maximizes the reward, which is sometimes referred to as being exploitative [81]. Similar to how infants explore their environment and discover the behaviors that enable them to accomplish a task, reinforcement models learn through interactions with the environment and observations to achieve a goal. In other words, reinforcement learning is trial-and-error learning based on rewards for desired behaviors and punishment for undesired behaviors.

All four types of learning approaches have played pivotal roles in advancing object detection, offering diverse avenues for developing accurate and efficient models for a variety of applications in computer vision. In the following, I will focus my review on (1) traditional methods for object detection, (2) supervised ML for object detection, and (3) unsupervised ML for object detection.

Traditional Methods

Conventional digital image processing algorithms focus on performing different operations on image data in order to get specific information and enhance visual information to facilitate human interpretation. These techniques involve the use of mathematical or statistical operations to modify images for many applications, including medical [82, 83] and satellite imagery [84], digital photography, object recognition and tracking, remote sensing,

augmented reality and industrial inspection [85]. Digital image processing involves various procedures, including data preprocessing, image enhancement and restoration, as well as image analysis to extract features, segment and classify specific data [86, 87].

Raw image data obtained from a camera does not often yield optimal results when used directly for computer vision tasks due to various factors such as environment noise and compression artifacts. Preprocessing is an essential step that encompasses the operations that are typically necessary prior to the main analysis of data and extraction of information. Some preprocessing techniques used to improve image processing results include noise reduction, image resizing, filtering and blur and focus corrections [88]. The selection of these techniques relies on the characteristics of the image and the intended application. Following are some well-known image processing techniques that are widely employed.

Image enhancement: Image enhancement refers to processing images with the aim of improving their usefulness by increasing the clarity and details. Depending on the application, the enhancement objective may be to improve perceptual aspects such as intelligibility and visual appearance- or to improve machine object detection performance. The algorithms used in image enhancement often exhibit simplicity and a qualitative nature [86]. An image can often be enhanced through various techniques including gray-scale modification (where the gray scale is change to a different gray scale based on a specific transformation), filtering (such as low-pass and high-pass filtering, median filtering and adaptive filtering), color and contrast enhancement (enhancing the colors and contrast between different elements within the image to make them more distinguishable), hue, intensity and saturation enhancement (an important color system to manipulate image appearance) and edge enhancement (to highlight and enhance the edges or boundaries between different objects or regions within an image) [89]. Image restoration and enhancement are closely interconnected processes. In the context of image restoration, the goal is to recover a degraded image and make the processed image closely resemble the original as much as possible. The degraded image is

assumed to be an imperfect version of the ideal or original image, and restoration techniques are employed to reduce or reverse the effects of degradation, such as noise, blurring, or other forms of distortion.

Image thresholding: Image thresholding is a popular technique used for image segmentation and classification, which involves separating an image into distinct regions or objects based on pixel intensity values. Thresholding techniques assign pixels to different classes or labels depending on various criteria. It is a commonly used technique for extracting foreground objects, creating binary images, or identifying specific regions of interest in an image. Common image thresholding algorithms include global thresholding, adaptive thresholding, and Otsu's thresholding [90]. In global thresholding, a single threshold value is applied to the entire image. All pixels with intensities above the threshold are assigned to one class, while pixels with intensities below the threshold are assigned to another class. This technique assumes a bimodal distribution of pixel intensities in the image. Adaptive thresholding is useful when the image has varying illumination or local intensity variations. Instead of using a single threshold for the entire image, adaptive thresholding calculates different thresholds for different image regions based on local characteristics. This allows for more accurate segmentation in images with uneven lighting or intensity gradients. Otsu's method is an automatic thresholding technique that determines the optimal threshold value based on the image's histogram. It aims to maximize the separability between foreground and background classes by minimizing the intra-class variance. This technique is effective when the image has a clear bimodal histogram.

Morphological operations: Morphological operations are a set of image processing techniques used to analyze and manipulate the shape, structure, and spatial features of objects within a digital image. These operations are based on mathematical morphology, which applies set theory concepts to image processing [91]. There are two fundamental morphological operations that are often used in combination to implement image processing

operations: erosion and dilation. Erosion removes pixels from the boundaries of objects in an image. It involves scanning the image with a small, predefined structuring element (often a small matrix or kernel) and replacing each pixel in the image with the minimum pixel value within the corresponding neighborhood defined by the structuring element. Erosion can be used to eliminate small noise or thin structures, shrink objects, and separate connected objects. In contrast to erosion, dilation adds pixels to the boundaries of objects in an image. Like erosion, it uses a structuring element and replaces each pixel with the maximum pixel value within its neighborhood. Dilation can be used to fill in gaps or holes in objects, enlarge or thicken structures, and merge nearby objects.

Traditional image segmentation and classification methods have been categorized into three following approaches as pixel-based, edge-based and region based methods [92].

Pixel-based methods: These methods also known as point-based segmentation, and encompass techniques such as thresholding in the feature space. The goal is to partition the image into meaningful regions or objects at the pixel level, by assigning labels to individual pixels in an image based on certain criteria or features.

Edge-based methods: These methods focus on detecting and delineating boundaries between different objects or regions based on intensity gradients or other edge-related properties. Common techniques include edge detection using operators like Sobel, Canny, or Laplacian, and subsequent edge linking or contour tracing [93].

Region-based methods: These methods divide an image into regions based on criteria such as homogeneity, texture, similarity and spatial proximity [94, 95]. These methods often involve iterative processes to merge or split regions to achieve desired segmentation. Examples include region growing, merging and splitting techniques and their combinations.

Traditional image processing methods have been foundational in the field of computer vision, primarily focusing on manipulating and enhancing images through a series of systematic techniques. These methods often involve preprocessing steps such as image

enhancement and image segmentation to improve the overall quality of the visual data.

Supervised Machine Learning

As technological progress continues, the integration of supervised machine learning methods seamlessly complements traditional computer vision techniques, ushering in a new era of enhanced capabilities in image analysis and interpretation. In supervised learning, a set of input data along with their known responses (outputs) is fed to the algorithm. During the training procedure, the algorithm tries to learn an effective way to map inputs to their corresponding outputs. Once the model is trained, it is used to predict the outputs for new data. In a supervised learning model, a preliminary step is to collect and label the data. Supervised learning consists of two main categories: classification and regression. Regression is the problem of predicting direct responses (such as numerical values), whereas classification has the goal of classifying the input into a specific category. To address classification task, there are many traditional supervised machine learning methods, such as Support Vector Machines (SVMs), Random Forest, k-Nearest-Neighbor, and Naive Bayes [96]. More recently, deep learning (DL) methods, including convolutional neural networks (CNNs) have been successfully used for classification and object detection. Traditional classifiers and object detectors are based on physically definable and extractable image features (shape, color, texture, and spectral features) that are manually chosen, whereas in DL methods, the selection of suitable image features is handled in the training procedure.

In 2001, the first real-time human face detection algorithm without limitations such as skin color was developed by Viola and Jones (VJ detector) [97, 98]. This detector incorporates integral image, feature selection and detection cascade techniques and utilizes sliding windows to go over all possible locations within an image and find the possible windows that contain a human face. A few years later, the histogram of oriented gradients (HOG) feature descriptor was introduced by Dalal and Triggs [99] using a linear SVM

detector. They used locally normalized histograms of gradient orientation features in a dense overlapping grid for person detection. The HOG detector scales the input image several times while maintaining the size of a detection window in order to detect objects of various sizes. The HOG detector has served as a crucial building block for numerous object detectors [100–102] and a wide range of computer vision applications [103]. As inspired by the HOG detector, the Deformable Part-Based Model (DPM) detector [100] achieved the best average precision in the 2006 PASCAL person detection challenge. The training procedure in this model can be thought of as simply learning the correct way to break down an object, and the inference can be thought of as an ensemble of various object parts. In this method, a margin-sensitive approach for data mining hard negative examples with a latent SVM is used to build a recognition system based on multiscale, deformable models. The fundamental building blocks for this model are the HOG features. Later on this model was enhanced to handle real world objects with greater variations [101, 104–106].

Rapid progressions in deep learning and improvements in device capabilities including computing power, memory capacity, power consumption, image sensor resolution, and optics have improved the performance and cost-effectiveness and further quickened the spread of vision-based applications. At present, deep learning-based object detection frameworks fall into two main categories: two-stage detectors and one-stage detectors. Two-stage detectors, consist of two stages: one stage to generate region of proposals as candidate boundary boxes and another stage to do object classification for each proposed region. In one-stage detectors, both stages are done directly on an image without proposing any pre-generated regions. Therefore, on public benchmarks, one-stage detectors are significantly more time-efficient and more suited to real-time applications, whereas two-stage detectors typically achieve higher accuracy [107]. Following, some representative methods in each group are described.

Two-Stage Detectors: In 2014, Girshick et al. [4] proposed a new object detection

algorithm by combining region proposals with CNNs. They called their method R-CNN, which stands for Regions with CNN features. Their object detection systems consists of three modules. The first module proposes regions which are category-independent, by using a selective search method. A CNN consisting of five convolutional layers and two fully connected layers is used in the second module to extract a 4096-dimensional feature vector from each region. To make the regions compatible with the CNN (inputs size of $227 * 227$), they are wrapped in a tight bounding box. The third module consists of a set of linear SVMs to classify each region. Fig. 2.5 shows this architecture. This object detection algorithm succeeded to give a 30% relative improvement over the best previous results on PASCAL Visual Object Classes Challenge 2012.

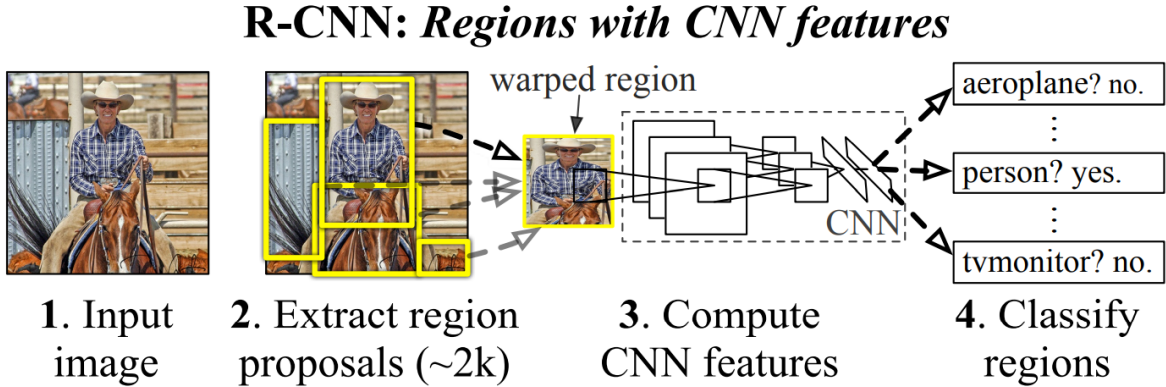


Figure 2.5: Architecture of R-CNN object detection [4].

Fast R-CNN [5] and Faster R-CNN [108], which are the fast and faster version of R-CNN, were introduced within one year after R-CNN. In R-CNN, a ConvNet forward pass is performed for each object proposal without sharing computation, hence it is slow on SVMs classification. Fast R-CNN uses pooling networks to speed up R-CNN by sharing computation. In this method, a convolutional feature map is computed for the entire input image and then each object proposal is classified using a feature vector extracted from the

shared feature map. In fact, both the full image and the identified regions of interests (ROIs) are fed into a CNN. Then each ROI is pooled into a feature map which is fixed in size, and then fully connected layers map them to a feature vector. Softmax probabilities and bounding-box regression offsets are two outputs for each ROI. This architecture is shown in Fig. 2.6.

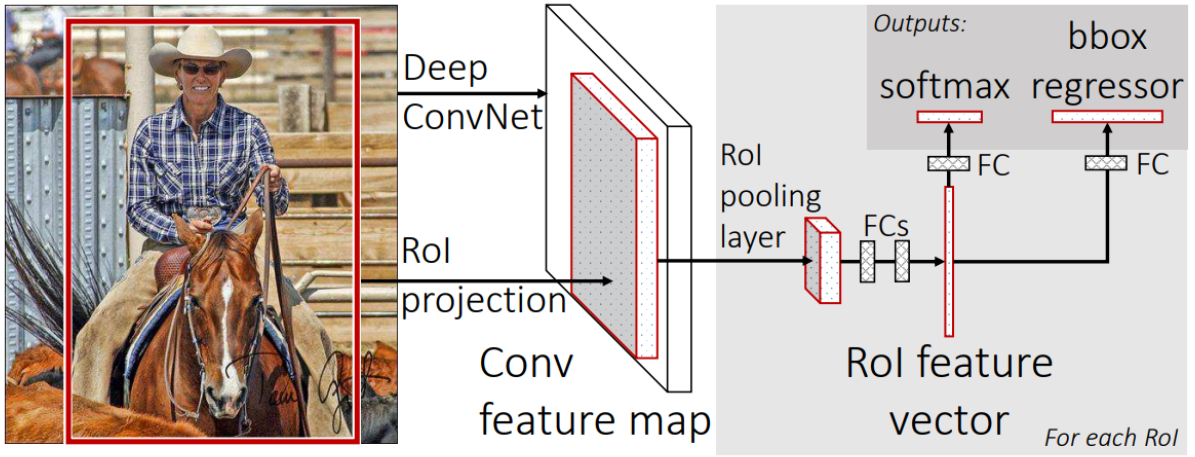


Figure 2.6: Fast R-CNN framework for object detection [5].

R-CNN and Fast R-CNN use selective search for proposal regions which is slow. Faster R-CNN, introduces the Region Proposal Network (RPN) to improve the region-based CNN baseline and enable nearly cost-free region proposals. RPN improves the baseline by sharing the full-image convolutional features with the detection network. It not only reduces the region proposal time, but also improves the detection accuracy. Up to now, all these methods were only for object detection and providing a boundary box around the object. Mask R-CNN [6] is a later extension of Faster R-CNN primarily for object instance segmentation. Each candidate object in Faster R-CNN has two outputs: a class label and a bounding-box offset. Mask R-CNN uses the same two-stage procedure with an identical RPN which is the

first stage as Faster R-CNN. In the second stage, by adding a branch for predicting an object mask, Mask R-CNN detects objects in an image while generating a segmentation mask for each instance. Mask R-CNN's framework for instance segmentation can be seen in Fig. 2.7.

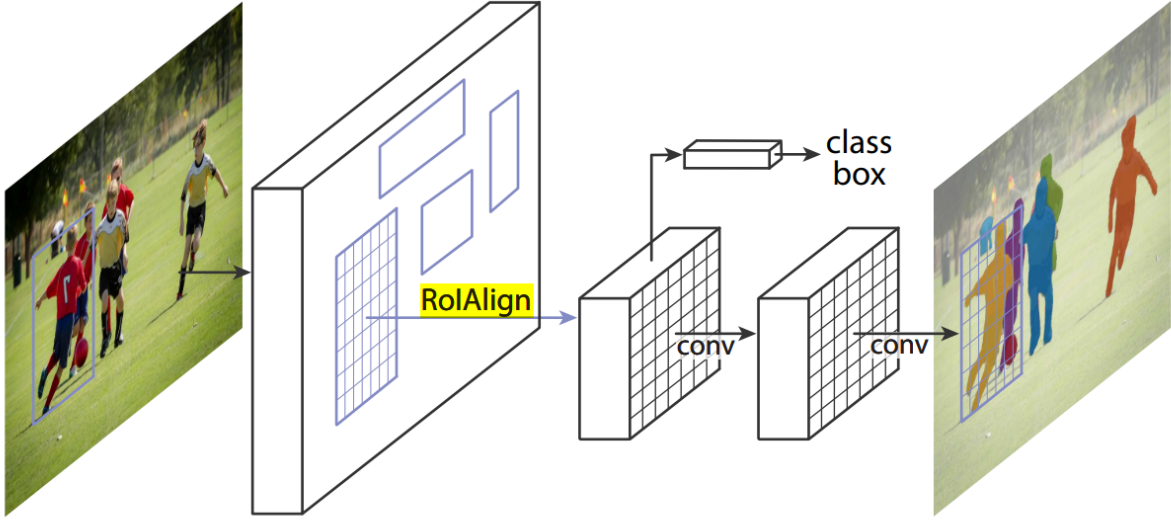


Figure 2.7: Architecture of Mask R-CNN framework [6].

One-Stage Detectors: You Only Look Once (YOLO) [7] is a well known one-stage object detector. In this algorithm, object detection is posed as a regression problem to spatially separate bounding boxes and corresponding class probabilities. Here, instead of two separate stages for proposing regions and classifying them after, bounding boxes and class probabilities are predicted only in one step and directly from full input images. Therefore, you only look once at an image to do object prediction. In this detection system, the input image is resized to 448 by 448, then it is fed to a single CNN which has 24 convolutional layers followed by 2 fully connected layers. Then, the resulting detections are threshold by the model's confidence score. This simple and straightforward detection system is extremely fast. The base YOLO model proposes 98 bounding boxes per image, which is much fewer than 2000

from selective search method. Since the first introduction of YOLO, it has been improved through time in terms of accuracy and speed to the various versions such as YOLOv2 [109], YOLOv3 [110] and YOLOv4 [111]. The YOLO detection framework is presented in Fig. 2.8.



Figure 2.8: YOLO framework [7].

Another famous single-stage detector is the Single-Shot Detector (SSD) [112]. In SSD, there is no region proposal generation and all computation is done using a single deep neural network. In this method, a feed-forward CNN produces a fixed-size collection of bounding boxes and allocates each box a confidence number. Final object detection is done using a non-maximum suppression step. In comparison to YOLO, SSD has extra feature layers at the end of a base network to predict the offsets to boxes with different scales, aspect ratios and confidence numbers. With an input image of size 300 by 300, SSD surpasses YOLO in terms of accuracy and speed.

RetinaNet [8] is another one-stage detector. This detector addresses the foreground-background class imbalance by proposing focal loss function as the classification loss that down-weights the loss assigned to well-classified examples. This focal loss concentrates on training hard examples rather than surpassing the training with immense negatives. Lin et al. have shown that RetinaNet outperforms previous one-stage detectors and two-stage detectors such as Faster R-CNN. This framework uses a Feature Pyramid Network (FPN) [113] as the backbone on top of a feedforward ResNet architecture [114] (Fig. 2.9).

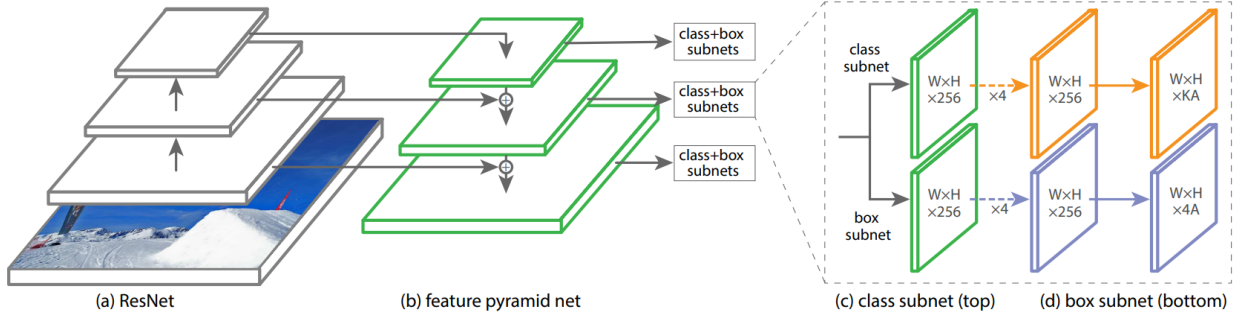


Figure 2.9: RetinaNet network architecture [8].

Unsupervised Learning

In contrast to supervised learning, unsupervised learning techniques try to automatically find structures and patterns in data without having it labeled. Computer vision techniques using unsupervised learning methods aim to extract meaningful information and patterns from visual data without the need for labeled examples or explicit supervision. Unsupervised learning approaches in computer vision often involve tasks such as clustering, dimensionality reduction, and generative modeling. Subsequently, I provide descriptions of some typical methods within each category.

Clustering: Clustering is one of the most common unsupervised learning techniques which is used to find hidden patterns in data and group inputs into clusters. Clustering methods can be categorized into two main groups of probability model-based approaches and nonparametric approaches based on a statistical point of view [115]. Probability model-based approaches assume that the data points are generated from a mixture model. Therefore, these approaches employ a clustering method based on the likelihood of the mixture model [116]. Nonparametric approaches in clustering primarily rely on an objective function that measures the similarity or dissimilarity between data points. These approaches can be further classified into hierarchical and partitional methods, with partitional methods being the most

commonly employed techniques [117, 118]. Expectation-maximization (EM) and K-means are two representative algorithms for the two groups respectively. The EM algorithm is an iterative process that aims to estimate the parameters of the mixture model and assign data points to their respective clusters. It consists of two main steps: the expectation step (E-step) and the maximization step (M-step). K-means partitions data into K clusters based on minimizing the distance between data points and cluster centroids. Clustering can be used for tasks like image segmentation and object recognition.

Dimensionality Reduction: Dimensionality reduction techniques aim to reduce the number of input features or variables while preserving the important structure and information of the data. These techniques are used to tackle the curse of dimensionality [119], improve computational efficiency, remove redundant or irrelevant features, and enable better visualization and understanding of high-dimensional data. Principal Components Analysis (PCA) [120, 121] is one of the most widely used dimensionality reduction techniques. It identifies the directions (principal components) in the data that capture the most significant variability. PCA transforms the original data into a new set of uncorrelated variables called principal components, where each component is a linear combination of the original features. By selecting a subset of the principal components, PCA allows for reducing the dimensionality of the data while preserving the maximum amount of information. Sparse coding is another dimensionality reduction technique that represents data as a sparse linear combination of a set of basis vectors. The training procedure attempts to find a sparse representation that minimizes the reconstruction error. By using a sparse representation, sparse coding effectively reduces the dimensionality of the data through selecting only the most relevant and informative basis vectors [122].

Generative Adversarial Networks: Generative Adversarial Networks (GANs) offer a means to acquire deep representations without the need for extensive annotated training data. GANs provide a framework through which meaningful representations can be learned,

leveraging the power of adversarial training instead of relying solely on labeled examples [123, 124]. GANs consist of two components: a generator network and a discriminator network, which are trained together in an adversarial manner. The generator’s task is to learn and capture the underlying distribution of real data and generating new samples accordingly. On the other hand, the discriminator serves often as a binary classifier, aiming to accurately distinguish between real samples and those created by the generator [125]. The learned representations from GANs have versatile applications across a range of tasks, such as image synthesis, semantic image editing, style transfer, image superresolution, and classification [124].

Applications of Object Detection Methods

Object detection is an integral part in a wide range of applications such as security systems, anomaly detection, robotics, healthcare and medicine, self-driving cars and face detection.

In recent years, autonomous vehicles have drawn a lot of interest as they have potential to benefit human lives by reducing traffic accidents, helping elderly and disabled people, and bringing comfort. The main key behind the advance of autonomous driving is real-time object detection models. These models enable the system to automatically identify and locate the objects around them such as pedestrians, vehicles and traffic signals. Many researchers have been working towards driverless technology, evaluating different object detection algorithms and the improvement of self-driving machinery [126–129].

Face detection, or identifying human faces in digital images and videos, is another example used in a variety of applications. Face detection is the backbone of all facial analysis algorithms, including those for face alignment, face modeling, face recognition, face authentication, head pose tracking, facial expression tracking, gender and age recognition, and much more, which enables human-computer interaction (HCI) [130]. Various work

has been done towards face detection using deep learning-based methods during the past few decades, showing the promising results of CNN object detectors in areas such as real-time detection and human face tracking [131], face detection [132–134], facial expression recognition [135–137] and multi-view face detection [138–140].

Security is another practical application area of object detection. The demand for effective security systems at airports is likely to be increased by rising crime. Dhiraj et al. [141] performed threat object detection by using two-stage and one-stage models to detect four classes of threat objects as gun, shuriken, razor-blade and knife. In another research, Khalid et al. [142] built a real-time monitoring system that is capable of detecting violent activities quickly and efficiently using deep learning models. This technology could benefit public areas by allowing security departments to take necessary actions to prevent crimes and violent activities.

The mentioned examples are just a few of the numerous applications of object detection that are under research and applied in every day life. With these promising results, I am highly confident that object detection will be an enabling technology for the mass production of 2D materials.

Time-series Forecasting

Time-series forecasting is an important applied machine learning technique that includes developing models to analyze (describe and predict) sequences of data collected during time intervals. These models are widely used in different areas such as finance [143], sales forecasting [144], climate time series [36], pore-water pressure [145], and medicine [146], to name a few. In time-series forecasting models, future values of a target $y_{i,t}$ for a given entity i at time t is predicted. The entity displays a logical grouping of temporal information. The most straightforward models which can predict the one-step-ahead can be represented as

below:

$$\hat{y}_{i,t+1} = f(y_{i,t-k:t}, x_{i,t-k:t}, s_i)$$

In this equation, $\hat{y}_{i,t+1}$ represents the model forecast, $y_{i,t-k:t}$ represent observations of the target for the previous k samples, $x_{i,t-k:t}$ represents observations of the input over the previous k samples, s_i is static meta data linked to the entity, and $f(\cdot)$ is the prediction model function which is learned by the model [41].

Recurrent Neural Networks

Deep learning algorithms have been developed and frequently used to extract information from many types of data during the last several years [147, 148, 148]. Recurrent neural networks (RNNs) are types of deep learning networks that can be used to analyze temporal information due to their specific architecture feedback loops which is their memory. They take information from prior inputs and update the current input and output [149] whereas in traditional neural networks, inputs and outputs are independent from each other. Fig. 2.10 illustrates the difference between a traditional feed-forward neural network and a recurrent neural network.

To be able to process the sequence data, RNNs have repeating modules arranged in chains (Fig. 2.11), where they can share the parameters across different time steps with the intention of employing these modules as a memory to store crucial data from earlier processing steps. The capacity of RNNs make them suitable for many applications such as natural language processing and language modeling [150, 151], speech recognition [152, 153] and emotion recognition in videos [154].

Over time, various kinds of RNNs have been developed and applied to temporal forecasting problems with strong results [155–158]. Recurrent connections can enhance neural network performance by taking use of their capacity for comprehending sequential dependencies. However, the techniques used for training RNNs might significantly restrict

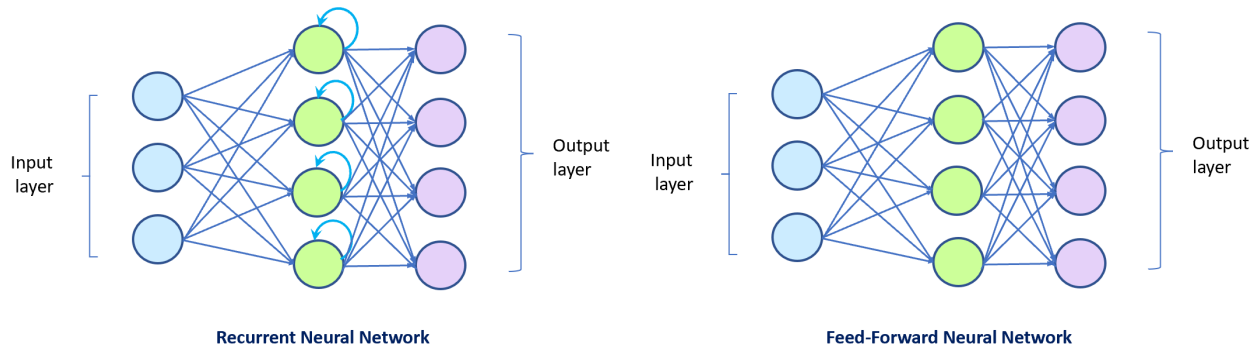


Figure 2.10: Recurrent neural network and feed-forward neural network architectures. The figure on the left shows a recurrent neural network with recurrent connections on hidden states which do the back propagation, whereas the figure on the right shows the feed-forward architecture of neural networks that have no recurrent connections.

the memory generated from the recurrent connections. Older variants of RNNs suffer from exploding or vanishing gradients during the training phase, causing the network to fail to learn long term sequential dependencies in data [159].

Explanation of LSTM Networks : Long short-term memory (LSTM) is the most popular method and model to address the insufficient memory of RNNs and tackling the problem of exploding and vanishing gradients. LSTM was introduced in 1997 for the first time to address the problem of long lasting learning to store information over extended time intervals by recurrent backpropagation [160]. In LSTM, instead of sigmoid or tanh activation functions, there are memory cells with inputs and outputs which are controlled by gates [160, 161]. In other words, LSTM is a modified version of RNN made of repeating modules called memory cells. These cells consist of three gates: an update or input gate, a forget gate, and an output gate. These gates work together handle learning long term dependencies. The input gate and output gate decide what information to update and pass on to the next cell respectively, while the forget gate decides the least relative information to throw away. This structure is shown in Fig. 2.12.

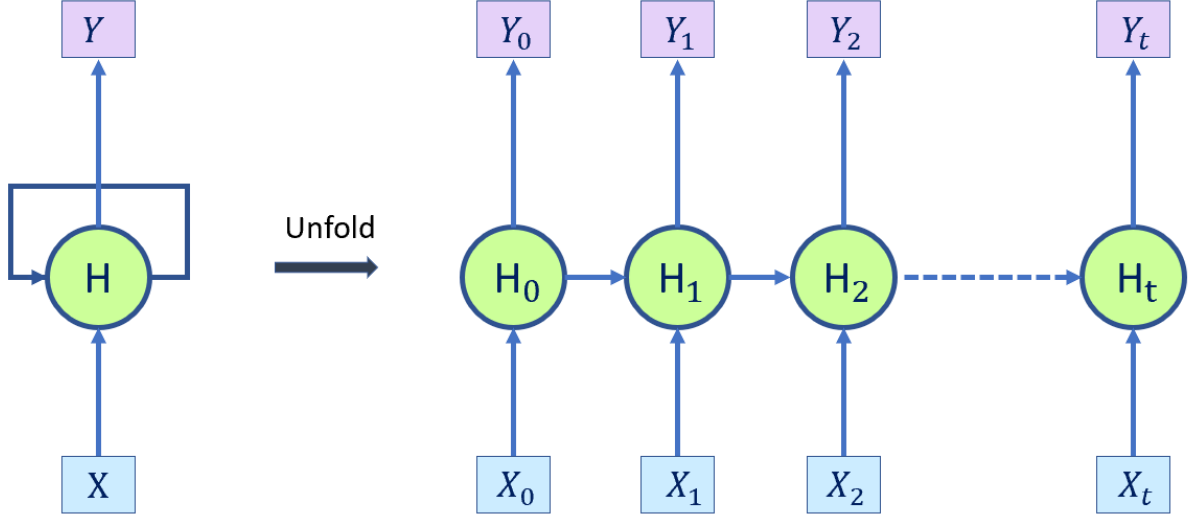


Figure 2.11: Unfolded sequential architecture in an RNN. X_t is the input at time step t , Y_t is the output at time step t and H_t is the hidden state at time t .

LSTM applications: LSTM neural networks can be applied to various tasks including prediction, pattern classification, recognition and analysis. Since LSTM is able to process sequential data, it is an effective tool in many different domains such as statistics, linguistics, medicine, transportation, computer science and more [162]. As instance, in [163] LSTM based models are used in natural language processing (NLP) for sequence tagging. In another work, Xue et al. [164] have applied LSTM to pedestrian trajectory prediction. They have proposed a network based on LSTM which is able to consider both influence of social neighbourhood and scene layouts in pedestrian trajectory prediction. Precipitation nowcasting is a crucial and challenging weather forecasting problem, in which the goal is to predict the future rainfall intensity in a local region over a relatively short period of time. Shi et al. [165] have used LSTM networks to build an end-to-end trainable model for the precipitation nowcasting problem. These instances are just a few of numerous applications of time-series forecasting models that are under research and applied in forecasting problems.

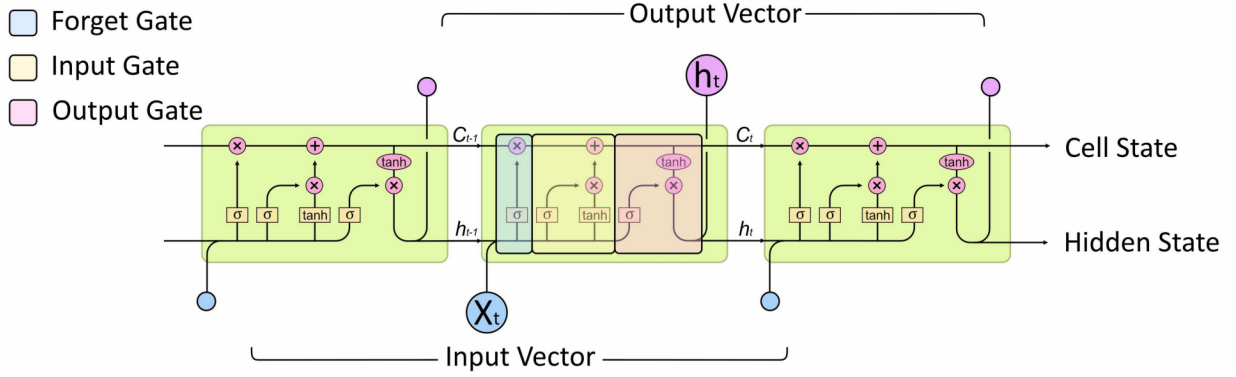


Figure 2.12: Memory cells structure in LSTM [9].

Achieving good and strong results in recent work was a motivation for me to apply time-series forecasting models to predict fluctuations in intensity and energy of QEs.

Correlation and Autocorrelation

Various physical phenomenon can cause variables within a dataset to be relevant or fluctuate in relation to each other. Finding and quantifying how reliant the variables in a data set are on one another or on themselves is crucial in many machine learning algorithms, including time-series forecasting problems. Correlation is a statistical measure that indicates this relationship. Correlation can be either positive, negative or zero. A positive correlation indicates that variables change in the same direction. A negative correlation indicates that the variables change in opposite directions, meaning that while one increases, the other decreases. A zero correlation indicates that the variables have no relationship with each other. The correlation coefficient is a unit-free statistical measure in the range of -1 and +1 that reports the correlation between variables. One of the most popular coefficient is Pearson's coefficient also known as the product moment correlation coefficient. It is the ratio between the covariance of two variables and the product of their standard deviations. Given a pair of variables (X, Y) , the formula for the coefficient represented by ρ is defined as below:

$$\rho = \frac{\text{cov}(X, Y)}{\sigma_x \sigma_y}$$

In this formulation, cov is the covariance, σ_x is the standard deviation of X and σ_y is the standard deviation of Y . For a sample and given paired data of X and Y , the Pearson's coefficient is commonly represented by r and can be expressed in terms of mean and expectation as below:

$$r = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^n (x_i - \bar{x})^2 (y_i - \bar{y})^2}}$$

where n is the sample size, x_i and y_i are the sample points indexed with i , and $\bar{x}$ and $\bar{y}$ are the sample means.

In time series problems, analyzing the correlation between a series and lagged versions of itself can provide predictive information. The correlation coefficient between a time series and its lagged versions over time intervals is called autocorrelation. In other words, autocorrelation determines the degree of similarity between a series and its lagged counterparts. Autocorrelation analysis is a popular tool for determining the degree of randomness in a data set. By measuring and analysing the autocorrelations for data values at various time lags, this randomness is determined. If the nature of the data over time is random, such autocorrelations ought to be close to zero for all time-lag separations. If the autocorrelation at one or more lag samples are significantly non-zero, then the data is not completely random [166].

2D MATERIAL DETECTION

ML for 2D Material Detection

Note: This chapter draws substantially from my published journal article: F. Ramezani, *et al.*, “Automatic detection of multilayer hexagonal boron nitride in optical images using deep learning-based computer vision,” *Scientific Reports*, vol. 13, p. 1595, 2023. This chapter also introduces additional topics, including the integration of the developed algorithm into a quantum materials pipeline.

Traditional Image Processing Methods

According to interference theory, the optical contrast between a substrate (typically a SiO₂/Si substrate with a known oxide thickness) and a 2D material varies based on the number of layers present. This optical contrast can be used to determine the number of layers in the material [167–170]. One of the first steps towards automating the detection of 2D materials dates back to 2011. Noelon et al. [171] proposed an automated approach based on an original image processing algorithm to determine the number of atomic planes in graphene samples. In their method first image of graphene sample is acquired via optical microscope. Then a set of image processing techniques is applied on the image to enhance it and identify the number of layers. These techniques includes micro-Raman calibration, light background subtraction, lighting nonuniformity correction, noise removal using median filter, and the color and gray scale image processing on each pixel of the image. This process result in a pseudo color map that identifies the graphene regions with different numbers of atomic layers as single-layer and few-layer.

In 2012, Li et al. [172] introduced a method to determine the layer number of 2D layered MoS₂ sheets on a 300 nm SiO₂ substrate using optical imaging together with an image analysis software called ImageJ. Using this software, the color optical image of the MoS₂

sheets is split into the monochromatic red (R), green (G), and blue (B) channels. By analyzing the high-contrast grayscale image of the MoS₂ sheets in the R channel, they determined the layer number of the sheets. The intensity difference between the MoS₂ sheets and 300 nm SiO₂ in the R channel increases with the layer number of MoS₂, allowing for identification of the layer number. To confirm the results obtained by the optical images, atomic force microscopy (AFM) and Raman spectroscopy were used.

In [173, 174], an optical method is purposed to identify the thickness of 2D nanosheets on transparent substrates using Boltzman function. In these works, they developed an optical technique to determine the thickness of 2D nanosheets on SiO₂/Si substrates by measuring the contrast difference between the nanosheets and the substrate. Later, the method was expanded to identify the thickness of nanosheets on transparent substrates, including polydimethylsiloxane (PDMS) film, cover slips, and glass slides. To measure the optical contrast values of the 2D nanosheets and the transparent, the ImageJ software was used. The optical contrast difference between the nanosheets and substrate was then derived by subtracting the individual optical contrasts. The values of all the measured contrasts differences were fitted with the Boltzmann function as a function of layer number.

To attain maximum reliability in identifying and analyzing exfoliated two-dimensional materials, it is crucial to prevent both false-positives and false-negatives in detections. In order to minimize the occurrence of false-positives and false-negatives, Jessen et al. [175] took various parts of the light path including the individual RGB spectral profiles of the light source, the 2D material interface, and the charge-coupled device (CCD) camera into the consideration. To identify the 2D crystals, they used the optical contrasts of 2D materials modeled in framework of multiple Fresnel reflections as an starting point. The contrast values obtained are used as a basis for a machine vision approach aimed at detecting the presence of 2D materials in digital images. Their machine vision algorithm first filter the image based on a specific range of RGB contrasts that corresponds to the expected values for the various

2D materials and their respective layer thicknesses that need to be identified. Then they use various sequential traditional image processing techniques (such as morphological operators and median filter) to reserve the large structures and remove the impluse noise, small and thin features, such as small particles, tape residues, and other types of contamination. Finally, the identified regions are highlighted using edge-detection algorithm.

The method presented in [176, 177], offers an approach for detecting whether a monolayer graphene is present or not in an image that contains n-layers of graphene material and other undesired particles using intensity based detection. Pixel intensities for different graphene flakes are affected by the thickness of the SiO₂/Si substrate, leading to obscuring effects [167]. For this study, 160 optical images were collected. Their approach involves utilizing a region of interest-based image segmentation process to remove irrelevant information from the image and isolate graphene particles. Afterwards, an intensity-based detection model that utilizes the unique color characteristics of monolayer graphene is implemented to distinguish it from other particles. This involved looking for differences in the color space of monloyarer as for red, green, and blue from the surrounding background pixels. These constraints help identify regions that contain monolayer graphene through a series of traditional image processing such as double thresholding and morphological operators. The output of the model identifies if the input image contains monolayer graphene and indicates the regions containing it.

Up to this point, some technological components including image analysis algorithms to identify flakes and layer-by-layer stacking techniques for CVD-grown 2D crystals were taken towards automation of 2D material detection [178]. However, a fully automated robotic production of vdW heterostructures that integrates vdW heterostructure fabrication, computer vision, and robotic technologies was not accomplished yet. As a step to achieve this goal, Masubuchu et al. [179] worked toward developing an autonomous robotic system to assemble vdW heterostructes, which they named 2D materials manufacturing system

(2DMMS). As part of fundamental prerequisites for any automated microscopy, 2D materials need to be detected. For this, they utilize optical images and applied various traditional image processing techniques including double filtering process based on the color contrast, information entropy (edge detection and morphological operations) in order to improve the earlier algorithm proposed in [171].

During the experiments of taking the optical microscopic images, uneven illumination from standard light sources and optical aberration could cause a significant image inhomogenities that can be an order of magnitude higher than the differences between layers. Therefore, Dickinson’s group [180] established a calibration to determine the thickness of a sheet by utilizing color/brightness that remains constant within an image and across multiple images taken under the same conditions. The optical imaging and image analysis procedure that they used comprised four major steps using Gwyddion [181] and ImageJ softwares: imaging, removal of inhomogenities, conversion of brightness to sheet thickness and recognition of individual sheets. For this, they used both low-magnification optical microscopy and AFM analysis.

Using traditional methods to detect 2D materials within an image is a time-consuming task that often take hours to accomplish. In another work on optical microscopic images, Lee et al [182] proposed an algorithm to reduce the time for searching the flakes. They used MATLAB as the image analysis algorithm to detect the number of layers through boundary detection. The accuracy of their method was assessed with the results achieved from conventional AFM and Raman measurements. Their algorithm can be divided into two main components: boundary detection and layer estimation. During boundary detection procedure, a boundary between the target flakes and the substrate is established. Then, it employs an adaptive RGB estimation process to compute the optical contrast between each target. This process generates a matrix with various integer numbers and provides a matrix map of them onto the target flake position to calculate the number of layers.

Optical microscopy is commonly used to determine the layer count of two-dimensional (2D) materials, but this method has limitations due to variations in illumination and camera settings across different labs. As a result, it is not possible to accurately determine the layer thickness of a material by comparing an optical image to standard or calculated images created under ideal conditions. Lei et al. [183], proposed an image reconstruction method to address this problem. Their method involves reconstructing raw optical images acquired with different illuminants and cameras into new images under specified illuminant and camera condition. This reconstruction process results in color differences that closely match those calculated under the specified conditions for each layer number. In this research color differences in the reconstructed image are compared to those computed under the given conditions, and the layer numbers of 2D materials in their lab and other research work are reported to be unambiguously identified.

In another research work in 2021 [184], Canny operator and edge connection iteration were used to seek out thin-layer WS₂ samples on complex surfaces. To determine the number of layers, three preprocessing procedures as demonising, smoothing and enhancement were adopted. Then the detection was done by edge detection and morphological operations followed by statistical based on a multispectral thresholding method to segment the regions of interest with different numbers of layers. To obtain the thresholding criteria, optical contrast was measured on a standard dataset of thin-layer samples of monochromic illumination photographs.

Supervised Learning Methods

In contrast to traditional computer vision methods, which involve lengthy trial-and-error processes, deep learning as a subset of supervised learning empowers end-to-end object detectors to attain higher levels of accuracy in tasks like image classification, semantic segmentation, and object detection. Deep learning algorithms have the ability to extract

high-level features from images through data training. As compared to traditional image processing algorithms, deep learning algorithms possess strong generalization capabilities and are resistant to variations in detection conditions. This characteristic enables them to avoid errors that can occur with traditional algorithms that require manual parameter adjustments.

As one of the first work toward 2D material detection using supervised learning methods, in 2018, Lin et al [31] developed a supervised machine learning optical identification (MOI) method based on SVM to enable identifications of the thickness, existence of impurities and tacking orders of graphene, molybdenum sulfide, and their heterostructures. Using optical microscopic images, their method is developed in two process: training and test process. During the training process, a comprehensive database of optical images of 2D materials was created to train the model and obtain the characteristic RGB information of the materials with different thickness. To do this, a set of 2D samples of photographs were taken at varying positions and light intensities. Each photograph was then segmented into sub-regions using the contour of the light intensity obtained from the substrate photograph to capture additional RGB information of the sample at different light intensities. Noise reduction was then achieved by applying median and mean filters to each image, followed by color calibration treatment that linearly scaled the G and B channels of all pixels in the entire photograph based on a manually selected substrate block. A manual mask was used for each photograph within its category to determine corresponding pixels (e.g., single layered graphene region), and the mean value of each channel with the same category in one sub-region was calculated as the training dataset. Finally, a 3D one vs. one linear SVM was used to train the model containing the characteristic RGB information using the training dataset. During the testing procedure, the RGB information of the images were collected and classified into specific categories. These classification resulted in a false-color images that indicates the distribution of the substrate, 2D material with various number of layers

and impurities. In another work, Yang and Yao [32] combined a SVM classification model with image processing techniques to identify 2D material flakes with different thicknesses, as well as measuring the flake’s sizes using color features in the images.

In one of the latest research deploying traditional supervised machine learning methods, Chao and his team [185] proposed a pixel-wise classification method to detect the layers of graphene based on tree-based algorithms. To decrease the number of training images and speed up the classification time, they considered decision tree, random forest, extreme gradient boosting (XGBoost) [186, 187] and light gradient boosting machine (lightGBM) [188].

To reduce the intensity dependency of optical images, Aleithan and his coauthor [189] worked on statistical and structural techniques used in image analysis that assess the interdependent relationships between pixels and link them to the characteristics of the underlying structure [190]. To do this, they analyzed the microscopic image texture of MoS₂ and WS₂ to distinguish monolayer from few-layer nanostructures. In their study, they employed grey level histogram (GLH), grey levels co-occurrence matrix (GLCOM) [191], and run-length matrix (RLM) [191–199] as three methods of texture analysis (TA) which correspond to first, second, and higher-order statistical methods, respectively. For each method, the Fisher coefficient [200] was used to automatically extract the features. As for classifiers, both traditional ML classifiers (linear discriminant analysis (LDA)) and neural networks (made of input and output layers + two hidden layers) were used. They reported an accuracy of 89 and 95 percent for MoS₂ and WS₂, respectively, using run-length matrix with neural network classifier.

As one of the first tries towards applying neural networks to identification of 2D materials, Greplova et al [201] developed three neural network consisting of 4 convolutional layers and one dense layer to detect the hBN flakes of thickness between 10 and 90 nm. As part of the preprocessing procedure, they divided the images with large standard deviation

into squares of size 80 by 80 pixels. Upon these preprocessings, they labeled the training set into two categories: "good" when the squared image contains a potential flake, and "bad" if the squared image does not contain any flakes. To augment the good samples, they used image rotation and mirror reflection techniques. These square images were then used as the input to the network for training. To address the imbalanced dataset of good and bad samples (as the good samples are very less in comparison to bad samples), they formed batches and trained them on the multiple defined networks. Through this iterating training, if a single network misclassify a relatively large number of bad flakes as good, it is unlikely that the other two models misclassify the same bad flake as good. By defining the accuracy as the percentage of correctly identified flakes withing the chosen evaluation batch, they reached the training accuracy between 87 and 88 percent, and reported a 55 percent of success on a new unseen dtata of 100 flakes.

By development of neural network architectures, Wu et al. [202] used SegNet network [203] as the the base of their deep learning algorithm. They used a dataset of 500 optical images collected and labeled in their lab. Having a test set of 100 images, they reported an average accuracy of 97.17 percent. In another work, Saito and his coworkers [204], designed a neural network based on U-Net [205] and trained it on a realistically small dataset of optical images (24 MoS2 and 30 graphene). They initially obtained 24 images of MoS2 and 30 images of graphene through an optical microscope, which they then augmented through various techniques such as random clipping, rotation, and color changes. They trained the network on each of the dataset individually to detect monolayer, bilayer, and other thicker parts (including trilayer, 4 layer and more than 5 layer). As a result, their network was able to distinguish between monolayer and bilayers with a success rate as well a cross-validation score of 70-80 percent on the test set of 11 images of MoS2 and 14 images of graphene. As observed in this method, the U-net has a failure to incorrectly identify the background (wafer) as a monolayer flake.

In optical microscopy, objects within images usually have high contrast in comparison to the background. Due to the limited features available in graphene microscopic images, there is a possibility of overlap between the pixel values of the background and graphene flakes, resulting in a high rate of false alarms at the pixel level [206]. In order to improve the false alarm rates and accuracy of graphene detection, U-Net architecture was modified by adding 5 convolutional layers in the decoder to generate more training parameters from the images [206]. Together with this modification, a white balance (WB) method to correct the color imbalance on the images, and a support vector machine (SVM) were added to segment the graphene flakes. This architecture was developed on a dataset of 246 images for training and 57 images for testing. This algorithm was able to achieve flake-detection rates of 87.09 for monolayer and 90.41 for bilayer graphene. Additionally, it was able to attain pixel-level accuracy of 99.27 for monolayer and 98.92 for bilayer graphene.

In another research done by Han et al. [30], they selected a deep learning segmentation algorithm based on VGG16 [207] for pixel wise identification of optical images of 2D materials. For this, they collected a total number of 817 images of 13 different 2D materials including graphene/graphite, hexagonal boron nitride (hBN), the 2H-phase semiconducting transition-metal dichalcogenides (TMDs) 2H-MoS₂, 2H-WS₂, 2H-WSe₂, and 2H-MoTe₂, the 2H-phase metallic TMDs 2H-TaS₂ and 2H-NbSe₂, the 1T-phase TMD 1T-HfSe₂, black phosphorous (BP), the metal trihalides CrI₃ and RuCl₃, and the quasi-1D crystal ZrTe as well as 100 images of only the background for training and test. They increased this number for training set to 3825 by augmenting the data. The trained algorithm is able to identify the material type as well as the thickness of individual flakes from one layer to six layers thick. Besides the type and thickness identification, their 2D material optical identification neural network (2DMOINet) can be used for the optical characterization of 2D materials and the extraction of deep graphical features from optical microscopy images which can be used to predict the material properties. They conducted a symmetric analysis to comprehend

how the neural network captures deep graphical features like color, contrast, edges, shapes, and flake sizes from optical images. With this analysis, the results obtained from the trained neural network can be employed to forecast the physical properties of materials.

With further improvement in deep learning based segmentation methods, Masubuchi et al. [3] implemented Mask-RCNN [6] and integrated it to an automated optical microscope to automatically search for 2D materials on SiO₂/Si substrates. The neural network architecture based on Mask-RCNN allows for the detection of exfoliated 2D materials and the generation of a segmentation mask and bounding boxes for each object. Using stochastic gradient decent method, the model is trained on a mixed dataset containing various 2D materials, such as graphene, h-BN, MoS₂, and WTe₂ in order to enhance the network's generalization ability. The dataset used in this study included 353 images of hBN, 862 images of graphene, 569 images of MoS₂, and 318 images of WTe₂. The number of annotated objects in each category were as follows: 456 for hBN, 4805 for graphene, 839 for MoS₂, and 1053 for WTe₂. For layer thickness classification, three categories were defined namely "mono" for one layer, "few" for two to ten layers, and "thick" for ten to forty layers. To improve the model performance, instead of training it from scratch, transfer learning was carried out twice: First transfer learning was applied on the model trained on MS-COCO dataset [208] and the model was trained on all the material categories to detect the material group as well as the layer thickness. Secondly, the previous weights were transferred as a start point to train the model on each of the subsets of the material individually. The findings of this two-step transfer learning suggest that training on multiple types of materials can enhance the accuracy of a model by allowing the backbone network to learn the shared characteristics of 2D crystals. At the end, they investigated the generalization capacity of the trained models on three different microscope settings. In deep learning based models, the network requires to be trained on a large amount of data labeled in order to achieve a considerable performance. However, in practice it is time consuming and costly to obtain and label a large amount

of images. To address the challenge of missing annotations in instance segmentation for the identification of 2D quantum materials, Xuan Back et al. [35] integrated Self-Attention and Soft-Labeling techniques into the Mask-RCNN. The experimental results applied on the same 2D material dataset in [3] demonstrated an out-performance in average precision by previous method by a significant margin.

To facilitate the information extraction for deep learning based models, Sanchez-Juarez's group [209] merged digital image processing with AlexNet [210]. As part of the digital image processing procedure, they used a series of filters to improve the image quality and maximize the contrast between the background and the monolayer. The filtered image was later merged with the the original image to create a well defined monolayer image as an input to the neural network model, allowing for a better identification of monolayers. The merge of two images results in contrast enhancement of monolayer and improves the visualization of them within the image.

To investigate the relationship between the network accuracy and optical contrast variations in MoS₂, a systematic study of different network architectures is done in [211] considering a common dataset. This study utilized three different deep learning methodologies: DenseNet [212] to classify, U-Net [205] to segment and Mask-RCNN to detect microscopic images of MoS₂ materials for automated atomic layer mapping. To evaluate the model performances, the three networks were trained and tested on various combinations of the original dataset as well as an unaugmented version of it, using two optical contrast variations: the same and different gamma contrast sampling indices per channel operations. Finally, the statistical analysis of the training and testing datasets, based on RGB optical contrast differences and CIE 1931 color space analysis, was employed to evaluate the performances of the neural network models.

Furthermore, the performance of U-Net in comparison to a modified neural network was carried out by Zhu et al [213]. To distinguish and characterize 2D materials and interface

distributions of 2D in-plane and vdW heterostructures, they developed a neural network with four hidden layers based on six primary color channels (RGB and HSV) of pixels of eight types of 2D materials taking monolayer and bilayer flakes. It is found that this modified network, outperforms U-Net in identifying interface characteristics in 2D heterostructure images due to its pixel-based identification protocol. The difference in color space information on each pixel between the images allows for this resolution. On the other hand, U-Net typically relies on capturing geometric shapes and edge features of various components in images as the primary means of image identification.

Based on the success of U-Net architecture in segmentation, later U-Net ++ [214] was introduced to improve the original architecture by redesigning the dense skip connections between the encoder and decoder, enabling each convolution block to learn semantic features from all previous blocks. Another related U-Net based model is the attention UNet [215], which utilizes a new attention gate (AG) model to focus on shapes and sizes of varying degrees. Based on these improvements in UNet, Yang et al. [216] proposed a two-level architecture including a modified UNet++ with AG and a ResNet classifier to identify and segment monolayer and bilayer graphene.

Scanning electron microscopy (SEM) images offer crucial information about synthesized materials, including the portion of the area covered by graphene, domain or grain size and shape, and nucleation density. These metrics are vital for evaluating the synthesized material quality. Graphene presents a visual contrast from its surroundings when viewed in an SEM, which makes it suitable for automated image segmentation [217]. Shah et al. [217] showcased the application of U-Net for automated segmentation of SEM images of graphene that was synthesized using chemical vapor deposition (CVD), anticipating that it could be utilized for measuring the cover of graphene through SEM analysis. During this procedure for training purpose, they take advantage of unsupervised techniques to generate the training set for U-Net to segment the SEM images of graphene.

In another work to use deep learning based models on graphene detection, Mahjoubi et al. [218] developed a pipeline to classify the exfoliated graphene flakes into six different categories based on thickness. They proposed a hierarchical framework to take advantage on both supervised and unsupervised learning for segmenting the flakes. To perform pixel-wise classification, the model was made up of k-means++ clustering with squared Euclidean distance and several convolutional neural networks. As part of this framework, they developed the segmentation model based on an object contextual representation network with a high-resolution network [219, 220]. Furthermore, a weak learning technique was introduced to help the neural network learn from new images while maintaining the knowledge acquired from previous images. This enables the neural networks to be retrained on a new data while preserving the previous knowledge. It is reported that this trained algorithm outperforms U-Net++ in terms of pixel accuracy.

Up to this point, most of the deep learning methods research have used high spatial resolution modality of microscopic images (eg. RGB images). To explore other imaging modalities such as hyperspectral imaging, Dong et al. [221] developed a 3D deep learning model that uses both high spatial resolution data (RGB images) and hyperspectral images (high spectral resolution) to identify and segment MoS₂ flakes into monolayer, bilayer, trilayer and multi-layer (more than 4). Hyperspectral imaging (HSI) is a method that expands beyond the traditional RGB (red, green, blue) color assignment of pixels and instead examines a broad spectrum of light. HSI breaks down the light striking each pixel into multiple spectral bands, allowing for a more comprehensive analysis of the imaged object or scene (cite to <https://www.sciencedirect.com/topics/medicine-and-dentistry/hyperspectral-imaging>). It is shown that the 3D deep learning model reaches an overall higher accuracy deploying high spectral resolution, in comparison to a 2D deep learning model (a single-stream U-Net (S-U-Net) model) trained only on RGB images.

Unsupervised Learning Methods

Unsupervised machine learning possesses the capability to uncover novel patterns in data. In contrast to supervised learning, it is often a more straightforward, faster, and cost-effective approach, as unsupervised learning eliminates the need for the manual labeling of data, which is a requirement in supervised learning. To demonstrate and simulate the visibility of 2D materials on a Si/SiO₂ substrate, Li et al .[33], adopted three features including optical contrast, total color difference and RGB, and utilized both supervised and unsupervised machine learning methods; k-nearest neighbors and k-mean clustering to measure the thickness of graphene and MoS₂ flakes.

In another unsupervised learning method, Masubuchi and Machida [34] used a series of traditional image processing techniques (such as Canny’s edge detection and morphological operations) on a large dataset of optical images of graphene on Si₂/Si substrates to extract the optical features. After feature extraction, to automate the identification of cluster centers and enhance classification accuracy, they used a clustering method based on Bayesian Gaussian mixture model with Dirichlet process (BGMM-DP) [222, 223] to detect the clusters into monolayer, bilayer, trilayer, 4layer and thick graphene flakes.

Instead of optical microscopy, in [224], the information obtained from atomic force microscopy topography and friction force microscopy (FFM) were merged to define the layer thickness of graphene. To analyze the data and define the graphene structures, spectral clustering [225] was employed which can be used for non-convex clusters and often surpasses conventional clustering methods like the k-means algorithm.

In a recent work towards applying unsupervised learning to 2D material thickness detection, Sterbentz et al [226] developed an unsupervised learning method including four clustering algorithms; mean shift [227–229], Density-Based Spatial Clustering of Applications with Noise (DBSCAN) [230, 231], multivariate Gaussian Mixture Model (GMM) [232–236] and Expectation-Maximization (EM). In their method, Mean shift and DBSCAN are

combined to determine the number of clusters as well as their centers in a given image. Once the centers are defined, GMM and EM are used to enhance the clustering as they allow fitting of data with arbitrary normal distributions. This enables the algorithms to handle materials and substrates that have different cluster distributions in RGB space rather than spherical clusters (despite of k-means clustering as it assumes spherical clusters).

Methodology

The reported convolutional network based models for 2D material detection, were developed on specific microscope settings under certain conditions such as image magnification, background color and brightness. Therefore, it limits them to be capable of using on other images acquired in other labs with different microscope settings. My proposed deep learning method to detect hBN flakes in 2D material microscopic images is shown as a pipeline in Fig. 3.1. This pipeline consists of three main steps which will be explained in detail.

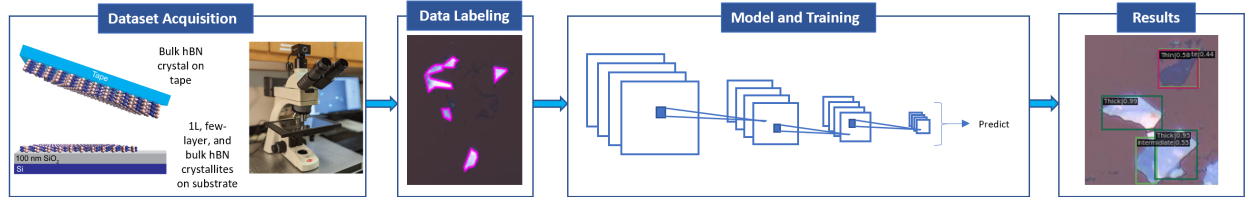


Figure 3.1: Proposed deep learning pipeline for 2D fake detection. First, a dataset of hBN samples are collected through mechanical exfoliation. Then as the second step, the dataset is annotated manually using an online annotation tool. During the third step, DetectoRS is chosen as the training algorithm, and for the last step, the trained model is applied to the test images to detect hBN samples.

Data Acquisition

To prepare the dataset, 208 images were collected over 2 months. To do this, hexagonal boron nitride (hBN) samples were fabricated via mechanical exfoliation. Bulk hBN crystals

were placed onto a piece of exfoliation tape and separated into thin layers by folding the tape together and peeling the tape apart five times. The exfoliation tape was then placed onto a $1 \times 1 \text{ cm}^2$ silicon (Si) wafer with a 100 nm silicon oxide (SiO_2) layer. To ensure the $1 \times 1 \text{ cm}^2$ region of the tape was in contact with the Si/ SiO_2 , the tape was pressed down with a cotton swab across the full area of contact. The sample, Si/ SiO_2 /exfoliation tape, was then heated for 1-2 minutes at 90°C and allowed to cool to room temperature, where the exfoliation tape was removed peeled. The hBN sample images were taken with a 20X objective using a Motic BA310MET-T Incident and Transmitted Metallurgical Microscope. The hBN images for the dataset were taken with the same camera settings and light intensity.

A second dataset of hBN images was collected by our colleagues at the University of Arkansas (UA) following a similar procedure. The UA dataset consists of 10 images that we use exclusively for testing the data. Finally, we use a third dataset for testing purposes: 10 hBN images collected by Masubuchi, et al. [3].

Data Labeling

Labeling the data is the preliminary step in a supervised learning object detection task, and the results of the model heavily depend on the accuracy of the annotations. To annotate our data, we used Roboflow (<https://roboflow.com>) which is an online annotation tool. For better accuracy of training, we annotated the flakes within each individual image manually and in three different categories as; Thick, Intermediate, and Thin. Table 3.1 shows the thickness and number of layers for each class. The table also indicates the number of images and the number of labeled instances of each flake type present in the full dataset.

In addition to annotating our images, we annotated the 10 test images in the UA dataset and re-annotated the 10 test images in the Masubuchi dataset to ensure uniformity in flake labels.

The annotations were then saved in the Microsoft common objects in context (COCO)

Table 3.1: Dataset Labels and Annotations

Label	Thickness (nm)	Number of atomic layers	Number of Annotations		
			Train	Validation	Test
Thick	≥ 50	≥ 100	996	240	73
Intermediate	25-50	50-100	465	14	12
Thin	≤ 25	≤ 50	140	23	7

JSON format. To train and evaluate the machine learning algorithm, the dataset was split into three subsets: training set, validation set and testing set. The training set consists of 177 images which contain 81% of the annotations; the validation set consists of 21 images which contain 14% of the annotations; and the testing set consists 10 images which contain 5% of the annotations.

Model and Training

To detect the 2D objects, DetectoRS [10] was chosen as the training algorithm. DetectoRS is a multi-stage supervised object detection algorithm. This algorithm involves two major components: Recursive Feature Pyramid (RFP) and Switchable Atrous Convolution (SAC). RFP is employed at the macro level that incorporates extra feedback connections from Feature Pyramid Networks into the bottom-up backbone layers. At the micro level, SAC is utilized, which convolves the features with different atrous rates and gathers the results using switch functions. Combining them results in DetectoRS, which significantly improves the performance of object detection. DetectoRS achieves state-of-the art performance on the COCO test-dev dataset. Fig. 3.2 shows this architecture [10].

The performance of supervised machine learning models depends on the quantity of the dataset to avoid overfitting. However, most of the time collecting and labelling the data is

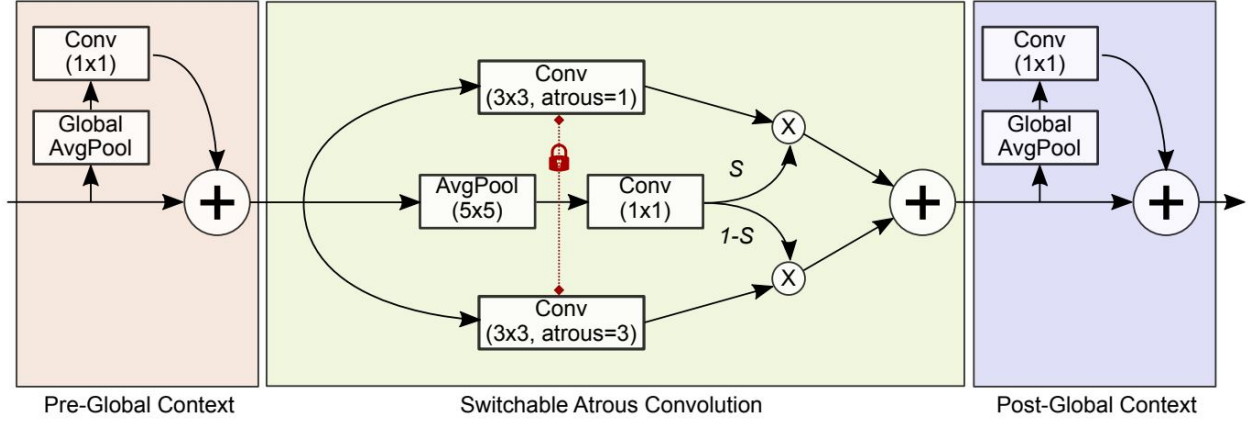


Figure 3.2: Architecture of Switchable Atrous Convolution (SAC) [10].

a tedious time-consuming process. Data augmentation techniques are a solution to address this problem. The augmented data will represent a more comprehensive set of possible data points, thus minimizing the distance between the training and validation set, as well as any future testing sets [237]. To augment the data, I used various techniques such as rotating, flipping and color contrast twice; once within Roboflow before exporting the annotation and another time during the training procedure.

To improve the performance of the model, I employed transfer learning. Fine-tuning pre-trained deep networks is a practical way of benefiting from the representation learned on a large database while having relatively few examples to train a model [238]. I took the model which was previously trained on the COCO dataset [208] and used it as a starting point to retrain it on our own dataset. For training on 2D material images, I chose DetectoRS algorithm with Cascade+ResNet-50 as the backbone detector architecture implemented in MMDetection library [239]. The model was trained with the learning algorithm rate of 0.0025, using Cuda and for 20 epochs.

Results

The inference results of the trained model when applied to one of the test images is shown in Fig. 3.3. Each flake identification in the image consists of three components; a bounding box, a class label, and a confidence number or probability score which shows the certainty of the algorithm that the class is detected correctly. As can be seen, the algorithm was able to detect the Thin sample, which is the most important among the three classes as it possibly contains a true monolayer material.

To measure the accuracy of the object detector across all test images, I used various evaluation metrics. Before examining these metrics, I present some basic concepts in object detection [240]:

- True Positive (TP): A correct detection of a ground-truth bounding box
- False Positive (FP): A detection of a nonexistent object or a misplaced detection of an existing object
- False Negative (FN): An undetected ground-truth bounding box

A correct or incorrect detection is determined based on Intersection over Union (IoU) which is a threshold. IoU measures how much of a projected object's bounding box overlaps with the bounding box surrounding the ground reference data.

It is also good to notice that a true negative result is when the algorithm detects the background correctly. Since labeled boundary boxes are only applied to objects, a true negative is not reported in the context of object detection.

Two common metrics to measure the accuracy of the object detection performance are precision and recall which are defined based on the TP, FP and FN: Precision = $TP/(TP+FP)$ and Recall = $TP/(TP+FN)$.

In words, precision returns the ratio of correctly identified objects to all identified objects. Recall measures the proportion of the ground-truth objects that were identified

by the algorithm. Mean Average Precision (mAP) is a measure that combines recall and precision for ranked retrieval results [241].

Another common tool used to measure the number (percentage) of correct and incorrect detections for each individual class based on the ground truth data is the confusion matrix. Fig. 3.4 shows how a confusion matrix is defined for four generic classes. Rows represent the ground truth for each class and columns show the predicted results.

In my research, I have considered each flake within an image for evaluation. Therefore I consider a true positive example to be when individual flake was detected and classified correctly. Incorrect detections occur when a flake is not detected at all (i.e. it is considered by the algorithm to be part of the background substrate), when it is detected as an incorrect class, or when the background substrate is detected as a flake. Fig. 3.5(a) shows the confusion matrix showing the classification results on the 10 images in the test dataset. It can be seen that my trained model was able to achieve a high number of correct detections, more notably for the Thin.

To assess the generalization of my method, I employed my trained algorithm using on ten images taken from two other datasets: The Masubuchi dataset and the UA dataset. I also compared my algorithm to the pre-trained algorithm published by Masubuchi et al. [3]

To make a fair comparison between my method and Masubuchi method, I fed these unseen images from both datasets to my algorithm. Later we manually annotated the same images and labeled them as ground truth. Then I calculated the confusion matrix for both datasets based on the ground truths and detected results. The obtained confusion matrices when tested on 10 images from each dataset are shown in Fig 3.5(b-c).

Fig. 3.6 and Fig. 3.7 show the detection results applying my proposed model to one image from the Masubuchi and UA datasets, respectively. In these images, my detector was able to identify 12 flakes correctly out of 16 annotated flakes within Masubuchi image, and 20 flakes correctly out of 32 annotated flakes for UA image.

I would like to note that the distinct color variations observed across the three subsets of 2D material images can be attributed to a combination of factors, primarily stemming from differences in microscope settings and variations in the substrates used for mechanical exfoliation. Microscope settings, including exposure time, white balance, and lighting conditions, contribute significantly to the color representation in optical images. Variations in these parameters across subsets can result in diverse color renditions. Additionally, the choice of different substrates for mechanical exfoliation introduces substrate-induced effects on material properties, reflections, and interactions, further influencing the observed colors. The combination of disparate microscope settings and substrate-dependent effects underscores the complexity of achieving consistent color representation across diverse datasets. Addressing these multifaceted challenges requires the establishment of standardized imaging protocols, calibration procedures, and collaborative efforts within the research community to enhance the reliability and comparability of findings in the study of 2D materials.

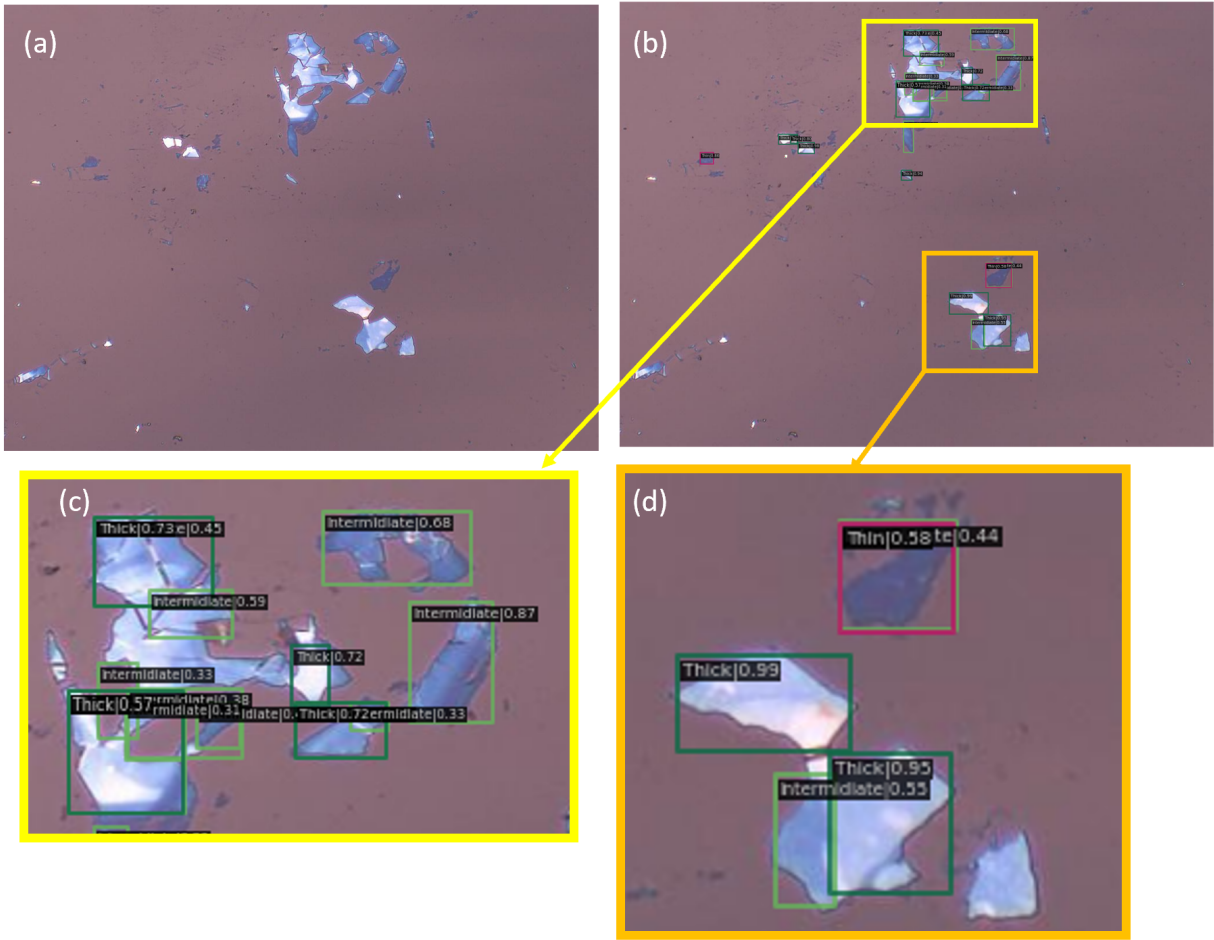


Figure 3.3: Results obtained from using the proposed algorithm on one image from our dataset; Part (a) shows the raw image, part (b) shows the labels generated by the algorithm, and parts (c) and (d) are zoomed-in regions showing the flake type and confidence score reported by the algorithm. We manually identified 19 flakes in the image, and the algorithm correctly classified 14 of them.

		Predicted Label				
		Class A	Class B	Class C	Class D	Recall
Ground Truth Label	Class A	TP _A	FP _B FN _A	FP _C FN _A	FP _D FN _A	$\frac{TP_A}{(TP_A+FN_A)}$
	Class B	FP _A FN _B	TP _B	FP _C FN _B	FP _D FN _B	$\frac{TP_B}{(TP_B+FN_B)}$
	Class C	FP _A FN _C	FP _B FN _C	TP _C	FP _D FN _C	$\frac{TP_C}{(TP_C+FN_C)}$
	Class D	FP _A FN _D	FP _B FN _D	FP _C FN _D	TP _D	$\frac{TP_D}{(TP_D+FN_D)}$
	Precision	$\frac{TP_A}{(TP_A+FP_A)}$	$\frac{TP_B}{(TP_B+FP_B)}$	$\frac{TP_C}{(TP_C+FP_C)}$	$\frac{TP_D}{(TP_D+FP_D)}$	

Figure 3.4: Confusion matrix template for a generic multi-class classification problem, showing how precision and recall are calculated for each class.

		(a) MSU Test Data				
		Predicted Label				
		Thick	Intermediate	Thin	Background	Recall
Ground Truth Label	Thick	53	2	0	18	0.73
	Intermediate	1	7	0	4	0.58
	Thin	0	1	3	3	0.43
	Background	20	13	4	0	0
	Precision	0.72	0.3	0.43	0	

		(b) Masubuchi Test Data				
		Predicted Label				
		Thick	Intermediate	Thin	Background	Recall
Ground Truth Label	Thick	47	14	0	4	0.72
	Intermediate	0	5	1	4	0.5
	Thin	0	1	5	4	0.5
	Background	17	26	6	0	0
	Precision	0.73	0.11	0.42	0	

		(c) UA Test Data				
		Predicted Label				
		Thick	Intermediate	Thin	Background	Recall
Ground Truth Label	Thick	95	6	0	66	0.57
	Intermediate	0	1	0	3	0.25
	Thin	0	1	0	9	0
	Background	94	40	0	0	0
	Precision	0.5	0.02	-	0	

Figure 3.5: Confusion matrix showing results obtained by applying the proposed algorithm to all three test datasets (MSU, Masubuchi, and UA).

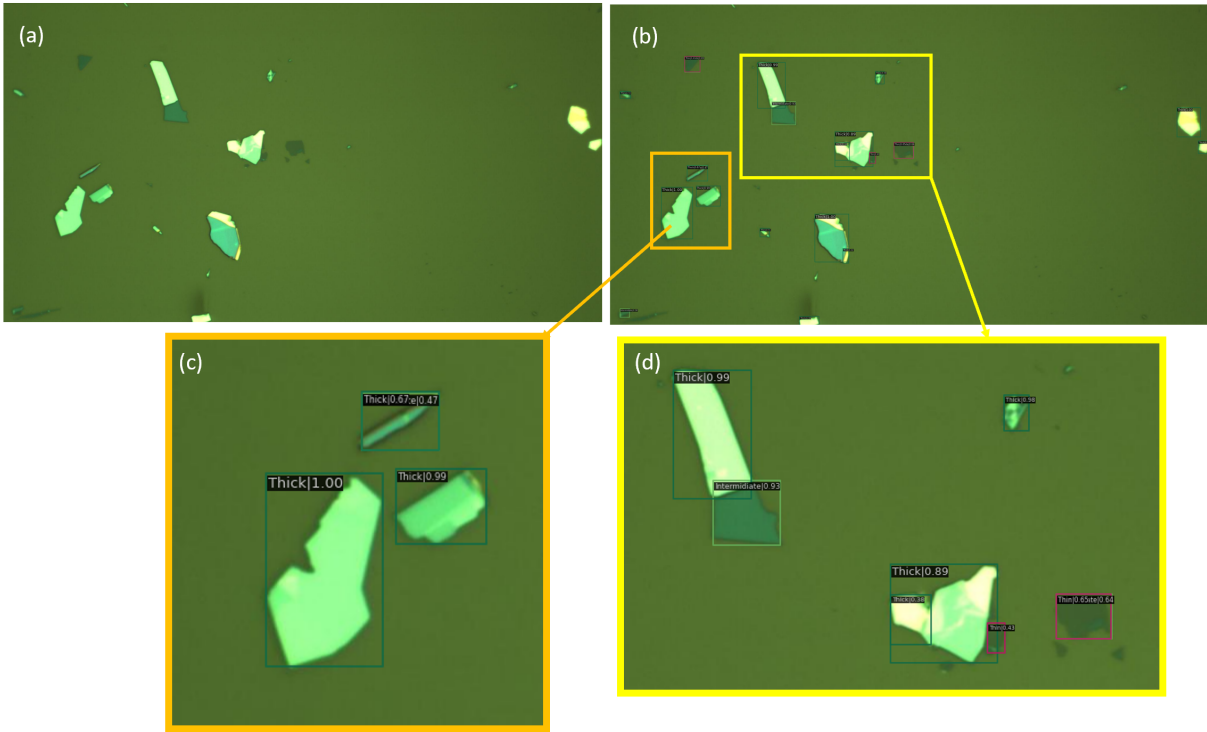


Figure 3.6: Results obtained from using the proposed algorithm on one image from the Masubuchi dataset; Part (a) shows the raw image, part (b) shows the labels generated by the algorithm, and parts (c) and (d) are zoomed-in regions showing the flake type and confidence score reported by the algorithm. We manually identified 16 flakes in the image, and the algorithm correctly classified 12 of them (microscope images obtained from <https://doi.org/10.6084/m9.figshare.11881053> under the license CC BY 4.0)

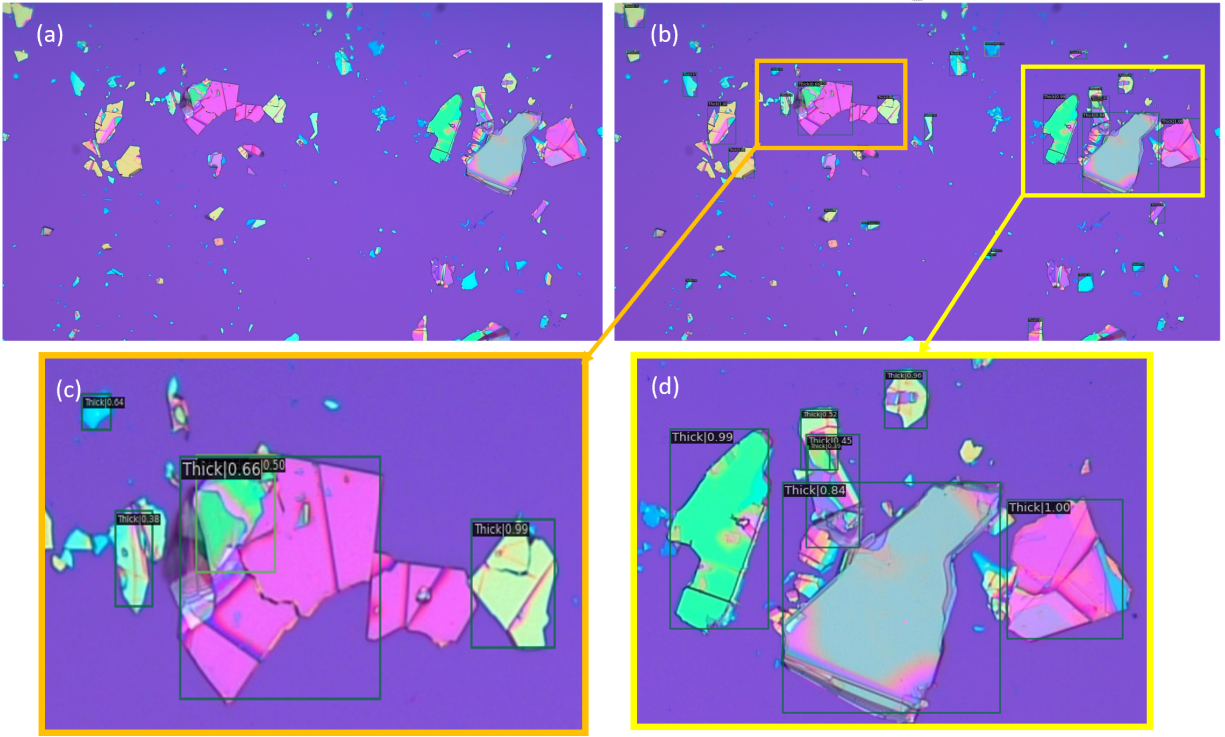


Figure 3.7: Results obtained from using the proposed algorithm on one image from the UA dataset; Part (a) shows the raw image, part (b) shows the labels generated by the algorithm, and parts (c) and (d) are zoomed-in regions showing the flake type and confidence score reported by the algorithm. We manually identified 32 flakes in the image, and the algorithm correctly classified 20 of them. (Microscope images obtained by Hugh Churchill and used with permission.)

		(a) MSU Test Data				
		Predicted Label				
		Thick	Intermediate	Thin	Background	Recall
Ground Truth Label	Thick	3	0	0	70	0.04
	Intermediate	0	0	0	12	0
	Thin	0	0	0	7	0
	Background	0	0	0	0	-
	Precision	1	-	-	0	

		(b) Masubuchi Test Data				
		Predicted Label				
		Thick	Intermediate	Thin	Background	Recall
Ground Truth Label	Thick	16	0	0	49	0.25
	Intermediate	1	0	0	9	0
	Thin	5	0	0	5	0
	Background	0	0	1	0	0
	Precision	0.73	-	0	0	

		(c) UA Test Data				
		Predicted Label				
		Thick	Intermediate	Thin	Background	Recall
Ground Truth Label	Thick	9	0	0	158	0.05
	Intermediate	0	0	0	4	0
	Thin	0	0	0	10	0
	Background	1	0	0	0	0
	Precision	0.9	-	-	0	

Figure 3.8: Confusion matrices showing results when applying Masubuchi’s pre-trained Mask R-CNN to our all three test datasets (MSU, Masubuchi, UA). The pre-trained model was unable to identify most manually labeled flakes in the test images.

Discussion

Before training the DetectoRS model, I applied Masubuchi’s pre-trained Mask-RCNN on our own dataset. I discovered this pre-trained model was not able to perform well on our dataset. For many test images, including the image shown in Fig. 3.3(a), the pre-trained model was unable to identify any flakes. For other images, the pre-trained RCNN was able to identify only a small subset of the total number of flakes. Complete results of using the Masubuchi method on all three dataset are presented as confusion matrices in Fig. 3.8. I note that one cause of the poor performance of their algorithm is that I modified the labeling procedure and criteria. Future work will explore a robust, standardized method for data labeling so that algorithm performance can be compared more directly.

In addition, future work will could include to collect and label data associated with additional 2D materials, expanding beyond hBN.

Object detection models, particularly those designed for identifying 2D materials with varying thicknesses, rely heavily on accurate annotations during the training phase. While

manual annotation serves as a crucial foundation for training deep learning models, it is essential to acknowledge the potential errors introduced during this process. Manual annotation, although a widely accepted practice, is inherently susceptible to human error. The process involves individuals visually inspecting images and delineating the boundaries of objects, in this case, 2D materials. Factors such as individual subjectivity, fatigue, and inherent biases can contribute to inconsistencies in labeling. One critical limitation in our dataset is the absence of quantum thickness verification for the manually labeled 2D materials. This lack of verification poses a challenge in establishing the ground truth with absolute certainty, leaving room for potential discrepancies between human-labeled data and true quantum thickness values.

Integrated 2D Flake Detection

Single and few-layers of two-dimensional (2D) materials have important fundamental properties allowing for their qualities in many potential applications, including many in quantum information science and engineering. In order to explore these materials for research and product development, single-layer and few-layer of 2D samples must be prepared through a manual exfoliation of a large population of smaller crystallites from a bulk crystal. This process result in crystallites with a different range of thicknesses which the entire population must be manually searched using optical reflected light microscopy to identify those that have a desired thickness (most often a single layer). This labor-intensive method of collecting single- and multi-layered samples severely hinders the rapid discovery and investigation of these material applications. The MonArk Quantum Foundry has been developing two prototypes of a 2D quantum material pipeline (2D-QMaP) to provide automated capabilities that unite and accelerate the three primary stages of sample preparation and device fabrication for 2D quantum materials research: 1) Robotic exfoliation, flake hunting, and stacking. 2) Device fabrication and packaging. 3) Characterization. As part of the

MonArk foundry, I have implemented a new deep-learning pipeline to detect and classify the crystallites of 2D materials based on coarse thickness classifications in reflected-light optical micrographs with a high detection accuracy. Using images with various microscopic settings, I have shown that my proposed pipeline is robust against changes in color or substrate background and is capable of being generalized to different systems. I have integrated my proposed pipeline to the prototype which has been under development at Montana State University, for a real-time 2D flake detection. This work, resulted in a novel fully automated optical microscope with ability to detect and classify 2D materials without the need of tedious manual searching through images which has never been done effectively. My proposed algorithm has enabled the 2D-QMaPs to automate the first tasks using a computer-controlled microscope. To achieve this ultimate goal, I have taken the following tasks as part of the procedure:

Set up consistent and sustainable labeling procedure Data labeling is the backbone of the computer vision supervised learning based developments. So far and for training, we have been using Roboflow, which is an online annotation platform with some limitation such as number of images and access of people. Therefore, I set up a consistent and sustainable labeling procedure such that an undergraduate student could effectively label and retrain the ML model for further 2D materials besides hBN. To do this, I worked with two free open-source annotation platforms; Make Sense and VGG Image Annotator (VIA). Make Sense (Fig. 3.9) is an open source web-based tool under GPLv3 license. Despite many annotation tools, Make Sense requires no advanced installation or operating systems, all you need to do is to open it through your browser. It supports multiple labels types such as rectangles, lines, points and polygons where is useful for accurate segmentation labeling. It also supports different output file formats besides COCO JSON such as YOLO, VOC XML and CSV that could be used for various object detection models. VGG Image Annotator (Fig. 3.10), is the second annotator tool which is simple and standalone manual software

that can be used to annotate various types of data such as image, audio and video. it runs in a web browser without needs for any prior installation or setup. The complete VIA software can be launched as an offline application which fits in a single self-contained HTML page of size less than 400 kilobytes.

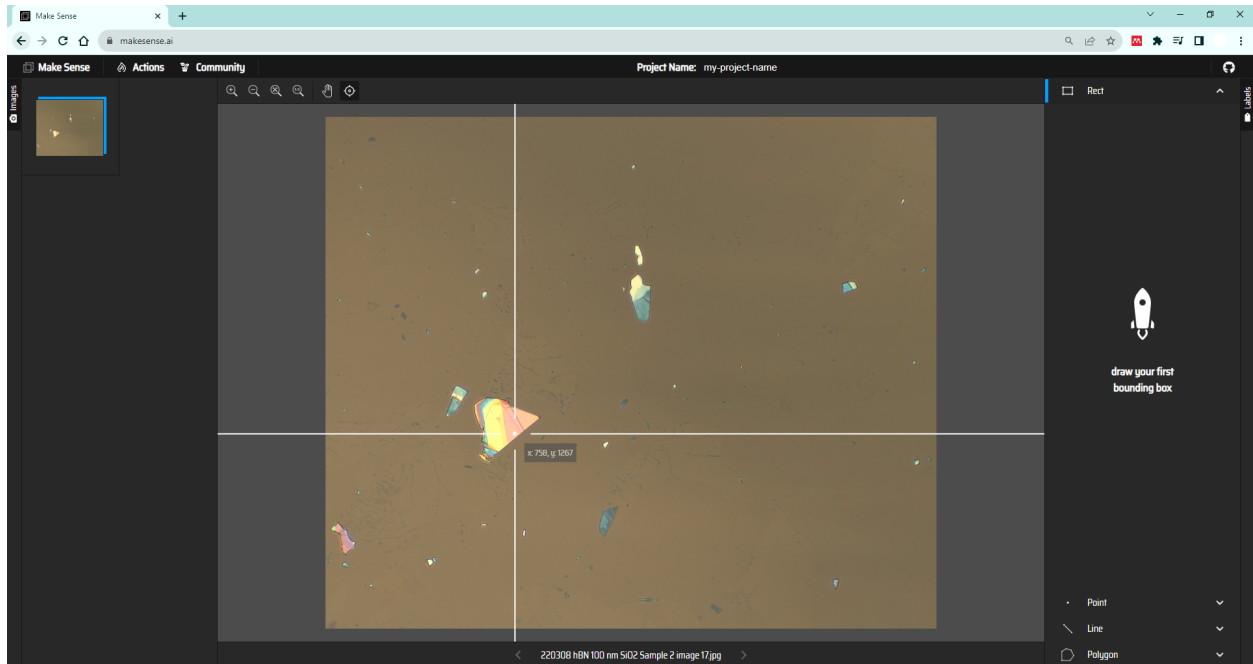


Figure 3.9: Makesense.ai is a free-to-use online tool for labeling photos

Migrate code from Google Colab to local PC The proposed model was trained and developed on Google Colab. Google Colab has three different plans as Colab, Colab Pro and Colab Pro+ which provide users with various access to GPUs and memory. The free colab plan, is the basic plan with the minimal access to the GPU and memory. The other two plans; Pro and Pro+, provide the user with faster GPUs and more memory. Despite of being one of the most popular platform for developing machine learning models in Python, Google Colab has a number of drawbacks. Pro and Pro+ limits GPU to NVIDIA P100 or T4, and RAM to 32 GB and 52GB respectively. Besides the limitation in GPU and RAM, their sessions are limited to 24 hours and Colab Pro does not support background execution.

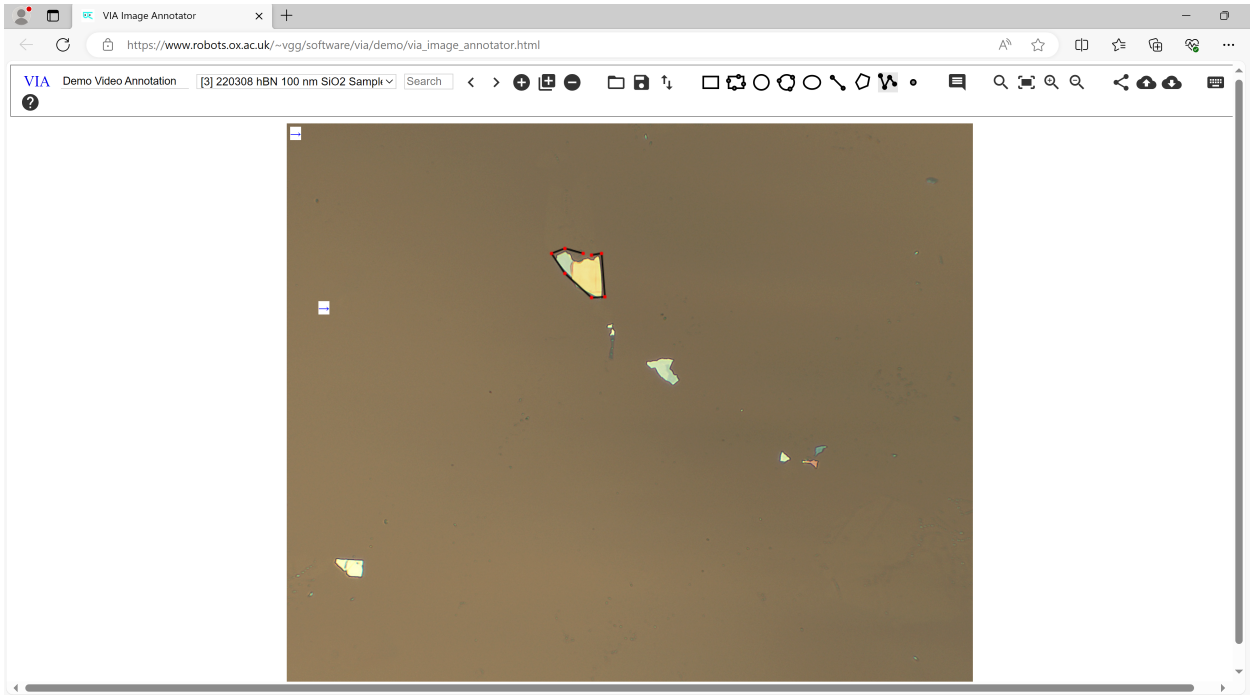


Figure 3.10: VGG Image Annotator is a simple and standalone manual annotation software for image, audio and video

They are also limited in runtimes which means every time you need to rerun and execute the training. Above all, none of the plans guarantee resources, so your instance may not be available at any time. These major drawbacks are not what is needed for an automated real-time procedure. To this end, I migrated the algorithm from Google Colab to a local PC with GPU sources and set up the necessary libraries. This independent the proposed real-time object detection model from external sources, and make it ideal for a local use.

Integrate proposed object detection algorithm to the microscope

As an achievement of my work, I have integrated my proposed 2D material detection method to the 2D quantum material pipeline (2D-QMaP) that has been under development by MonArk Quantum Foundry to provide automated capabilities that unite and accelerate the primary stages of sample preparation and device fabrication for 2D quantum materials

research. My suggested algorithm has empowered 2D-QMaPs to achieve a completely automated and real-time detection of 2D flakes, a task that was previously not effectively accomplished. To this end, I used ScopeFoundry which is an open-source local Python library. It is a platform for controlling custom laboratory experiments and visualizing scientific data. I integrated the developed object detection algorithm to this platform. As a result, acquired images can be processed with the proposed detection algorithm through an end-to-end automatic procedure. ScopeFoundry is a an open-source alternative to proprietary instrument control softwares like LabVIEW or MATLAB. What sets ScopeFoundry apart is its modular, cross-platform Python graphical interface, which facilitates rapid data acquisition and visualization, which empowers users to construct graphical interfaces for their lab equipment interactively using Qt Creator. This user-friendly approach not only streamlines the development process but also offers live updates of measurement code, expediting debugging and fine-tuning. Its also provides Hardware plug-ins for simple and complex scientific equipment. Currently, ScopeFoundry is used in multi-modal scanning microscopy measurements with electrons and optics with flexibility to be extended to many other data acquisition tasks. Fig.3.11 shows the results of the integration of my proposed 2D material detection algorithm to the optical cataloger.

Make the data and algorithm publicly available To help the other researchers and scientists explore improved methods for solving 2D material object detection, the algorithm as well as the data is published publicly under the title of "Automatic detection of multilayer hexagonal boron nitride in optical images using deep learning-based computer vision" at Scientific Reports with the DOI 10.1038/s41598-023-28664-3 and are available on GitHub at http://github.com/BMW-lab-MSU/hBN_Detection.

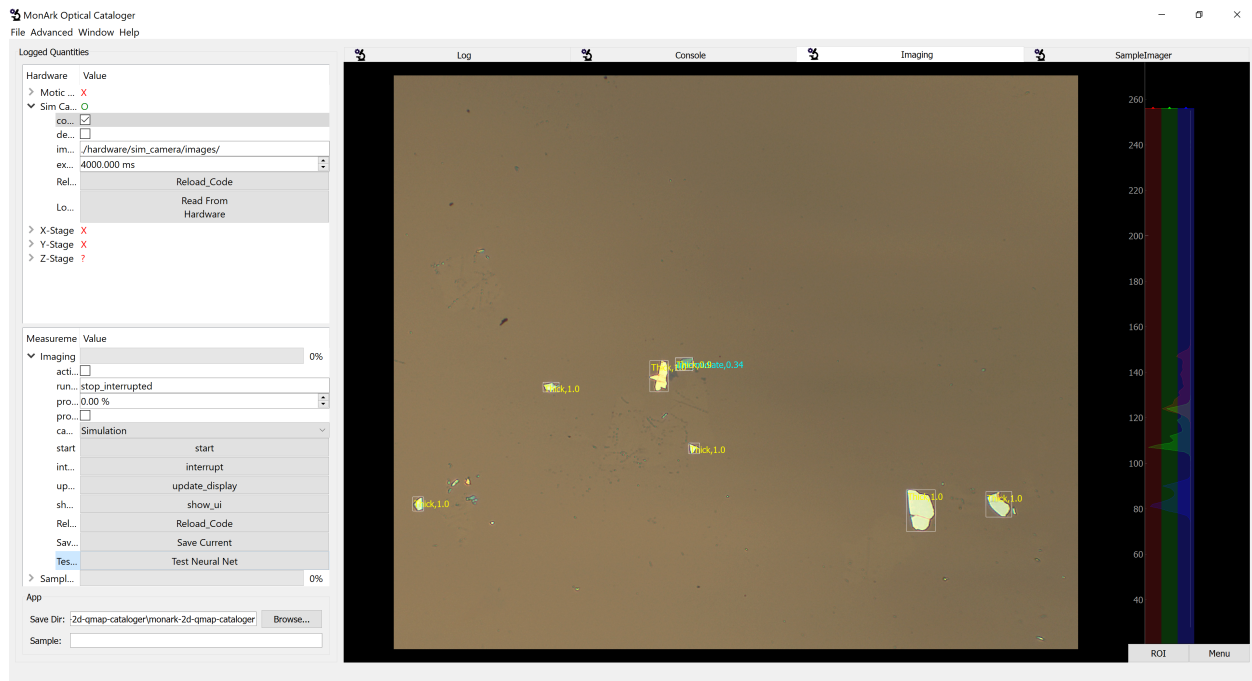


Figure 3.11: Real-time 2D material detection

QUANTUM EMISSION PREDICTION

Note: This chapter draws substantially from my submitted journal article: F. Ramezani, *et al.*, “Predicting Quantum Emitter Fluctuations with Time-series Forecasting Models,” *Scientific Reports* (submitted 2023).

Real world applications of quantum technologies require a novel manufacture which is able to implementing quantum algorithms and information. Photonic quantum systems provide implementation potentials for all major areas of modern quantum technology, including quantum communication, quantum sensing, quantum computing and simulation [68]. Integrated quantum photonic circuits are a desirable technology platform due to their strong potential to reach the level of component integration and performance needed for information processing [242]. To maximize the usefulness and value of integrated quantum photonics, the wavelength, polarization and intensity of photons within the QE should be stable. However, harmful fluctuations in the QE caused by local environment is unavoidable. A solution to this problem would be developing the QE devices that are dynamically tunable and apply a corrective feedback system. As part of the MonArk Quantum Foundary, in parallel with development of the QE devices and acquiring dataset of time-dependent QE fluctuations from the PL of nanobubbles of 1L-WS₂ and 1L-WSe₂ as a function of temperature, I assessed the random nature of the quantum fluctuations, and presented time series forecasting deep learning models to analyze and predict QE fluctuations for the first time. My trained models can roughly follow the actual trend of the data and, under certain data processing conditions, can predict peaks and dips of the fluctuations. The results of this algorithm provides the physicians with necessary information to design possible actions to tune the QE dynamically, as external stimuli affects emission spectra [73, 243, 244]. This idea of predicting quantum fluctuations is very novel, as currently no one is able to anticipate when they occur. The ability to anticipate these fluctuations will

allow physicists to harness quantum fluctuation characteristics to develop novel scientific advances in quantum computing that will greatly benefit quantum technologies.

Quantum phenomena such as superposition and entanglement offer many opportunities to revolutionize secure communication [64], computation [65], simulation [66], and sensing [245] technologies. Realization of these opportunities necessitates the development of new arsenals of material, devices, and control routines to robustly generate quantum states (qubits) and perform quantum operations (gates). In particular, the field of quantum photonics focuses on using single photons of light as qubits to both transmit quantum information and perform quantum processing operations. Single photons are resilient to decoherence effects, making them ideal quantum information carriers. Thus classical integrated photonic technologies offer an exciting foundation on which to build and deploy wafer-scale quantum photonic systems [68]. Generation of single-photon states is a fundamental requirement for quantum photonics, where the ideal solution has a small footprint to facilitate incorporation in integrated photonic circuitry. Additionally, single-photon states will enable the production of photons at GHz rates, electrical triggers for on-demand photon generation, and indistinguishable photons with identical polarization states and wavelengths [246].

Generation of single photons of light with solid-state materials is an attractive solution for quantum light sources for integrated photonics [67]. Unlike solutions that rely on nonlinear optical processes, solid state quantum emitters (QEs) can be triggered on-demand and have the potential for miniaturization for incorporation into integrated photonics. In 2015, solid-state QEs were discovered in the two-dimensional (2D) semiconductor single-layer WSe_2 (1L- WSe_2) [247–251]. While this new class of quantum light sources has many exciting properties, the wavelength (and presumably the polarization) of the photons emitted significantly fluctuates on the timescales of seconds (i.e. spectrally diffuses/drifts/wanders), significantly diminishing their performance in terms of photon indistinguishability [252].

From a materials/devices perspective, increased stability may be achievable by increasing the strain in the system [244], tuning the charge density in the material [253], and/or modifying the surrounding dielectric environment [254, 255]. Additionally, because the QEs in 2D materials are responsive to external stimuli [256], the emitters have the potential to be actively monitored and stabilized with a feedback loop during operation.

Here, I present an initial exploration of the ability for an neural network-based machine learning algorithm to predict the fluctuation of QEs in 2D materials based on the immediate history of its emission. Such a predictive algorithm could be used to improve photon indistinguishability of an advanced QE device where the emission wavelength is monitored and external stimuli is applied to prevent predicted fluctuations. My results show that a neural network that is rudimentary trained on the intensity fluctuations of discrete wavelength bins may be able to forecast future fluctuations. The work sets the stage for more sophisticated training strategies that take into account both emitter intensity and peak emission wavelengths.

Fluctuations of QEs in 1L-WS₂

Fig. 4.1 shows representative deleterious fluctuations of QEs from randomly occurring nanobubbles [257] in single-layer WS₂ that is deposited on a gold surface with gold-assisted exfoliation techniques [258, 259]. The QEs are identified as narrowband emission lines that are superimposed on a weaker background of broader, excitonic states. Inspection of the time evolution reveals that the center wavelength and the intensity of individual states changes on the timescale of seconds. These fluctuations are likely caused by the local environment surrounding the emitter, including nearby metallic and dielectric structures. The quantum emissions occur in four wavelength bands, approximately centered at 606 nm, 613 nm, 621 nm, and 629 nm. Large-area encapsulation of the devices with thin layers of hBN provides potential means of mitigating these fluctuations. However, they are not an

absolute solution to eliminate them completely. To address further stability, QEs that are dynamically tunable can be used alongside a feedback system that intermittently samples QLED emission and applies corrective stimuli. Machine learning time-series forecasting models have the potential to predict QE fluctuations and provide physicists with essential information to apply corrective stimuli to stabilize their emission wavelength and intensity in real time.

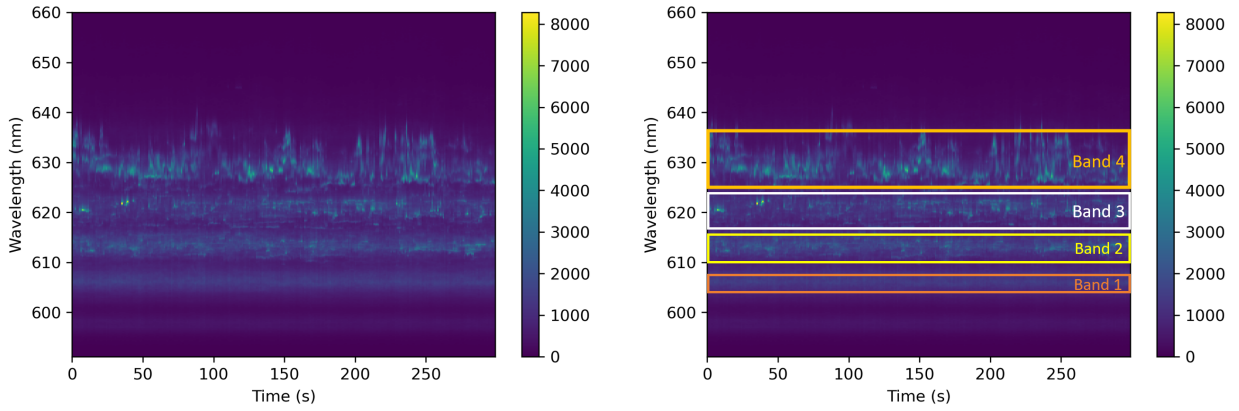


Figure 4.1: Fluctuations in intensity and energy for a nanobubble QE in single-layer WS_2 on an Au surface. The figure on the left shows the number of photons emitted by the device as a function of time and photon wavelength. The figure on the right shows the same data, but also indicates four different wavelength bands where quantum emissions occur.

Methodology

Data Acquisition

To measure the luminescence of a single bubble, we use a standard confocal fluorescence microscope in conjunction with a closed-cycle liquid helium cryostat with optical access (Montana Instruments). The sample is cooled to <4 K to slow the ionization of the localized states due to heat. A focused laser beam (wavelength = 532 nm, continuous wave, FWHM <1 μm) is then directed onto an area that contains narrowband emitters with energies

less than the exciton state of WS_2 . The power of the laser is adjusted to be below the saturation threshold of the emitters. The photoluminescence from the sample is collected and spectrally isolated from the reflected laser light with a series of thin-film interference filters. The fluorescence is then spectrally dispersed with a Czerny-Turner optical spectrometer (Princeton Instruments) and measured with a cooled scientific CCD camera (Princeton Instruments). The continuous series of spectra are acquired from the region of the sample for 180 s with 2 s integrations per spectrum.

Preprocessing

In order to learn historical values and patterns to forecast the future, there needs to be a window of time that is fed to the algorithm, so it can make predictions based on the previous values. For this, I transformed the data into input and output pairs where the observations at 5 previous time steps (a window size of 10 seconds) are used as inputs to make predictions for one step into the future (2 seconds). To later train and evaluate the forecast models, I divided this dataset into three subsets: a training set, a validation set and a test set. The training set consists of 80% of the time sequence, the validation and the test sets each consist of 10% of the time sequence of the data.

The performance of machine learning models depends on the quality of the data. There are various preprocessing techniques which convert the data into useful information to the models. Among various techniques, data normalization is an approach which entails scaling or transforming the data to ensure that each feature contributes equally and avoid bias towards a specific range of values in the feature data. As part of the preprocessing of my algorithm, I first checked the dataset for negative photons counts and zeroed them out as they are produced due to noise and/or calibration error. Then I normalized the data based on the statistical mean and standard deviation of the training set. Each value of the data (including those in the validation and test sets) is converted as follows, where X_{new} is the

normalized value, M is the statistical mean of the training set and σ is the standard deviation of the training set:

$$X_{new} = \frac{X - M}{\sigma}.$$

Experiment and Model

To forecast the photon intensities, I designed two different scenarios:

1. Prediction considering all measured wavelengths
2. Prediction considering 4 different bands including the most correlated wavelengths

For the prediction models, I developed a multivariate LSTM model and a shifted forecast that considers the current measurement as a prediction for the next step. As there are more than one variable in the dataset, a multivariate time-series forecast can be used. This will consist of multiple variables, each of which is dependent on other variables to some extent, in addition to its own historical values. For the first scenario, I developed a model with inputs of 5 time steps (10 seconds) over all the wavelengths regardless of their correlations, and then I did predictions for the values one-single step into the future. For the second scenario, I divided the wavelengths into the 4 bands that include the most correlated wavelengths. (See Fig.4.1.) Then I trained the model on each band individually considering the same time steps as the first experiment and predicted one-step in the future.

In my research I used various open libraries in Python, such as Tensorflow, Matplotlib, Pandas, Numpy for preprocessing, training and visualisation purposes. My LSTM model consists of one LSTM layer followed by two fully connected layers to predict the feature values. I employed a sequence length of 10 seconds as inputs, and predicted 2 seconds in the future. The LSTM models were trained using Adam optimizer with a learning rate of 0.0001 and a mean squared error as the loss function. Furthermore, during the training, I employed

callbacks to prevent the overfitting of the network and saving the best performance results based on the error obtained on the validation set.

Evaluation

To evaluate the models, I used root mean square error (RMSE) as the cost function which is commonly used in forecasting problems [260–262]. Considering n as the number of samples, y_i the actual value and $\hat{y}_i$ predicted value, RMSE is defined as:

$$\text{RMSE} = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2}.$$

For each of the experiments and models, I report the RMSE. I also present plots of the actual intensity values and predicted values, as well as the differences between these two.

Results

Correlation

To divide the data to bands with the highest correlations, I measured the correlation matrix and picked the wavelengths with the highest correlation coefficient using Pearson’s algorithm. This resulted in 4 bands with different numbers of wavelengths, as described in Table 4.1. The correlation coefficient matrices are shown in Fig. 4.2, Fig. 4.3, Fig. 4.4, Fig. 4.5 on the left. Before starting to train the time-series forecasting models, I assessed the randomness of the data to ensure the problem is predictable. I measured the autocorrelation for the dataset for 30 lags (60 seconds) and plotted the center of each band for illustration purpose in Fig. 4.2, Fig. 4.3, Fig. 4.4, Fig. 4.5 in the right. The blue shaded area is the confidence interval and represents the significance threshold. This indicates that autocorrelations inside the blue area are statistically insignificant, and anything outside of it is statistically significant. It is visible that, for each of the wavelengths, there are several

autocorrelations that are significantly bigger than zero, therefore the time series dataset is not completely random.

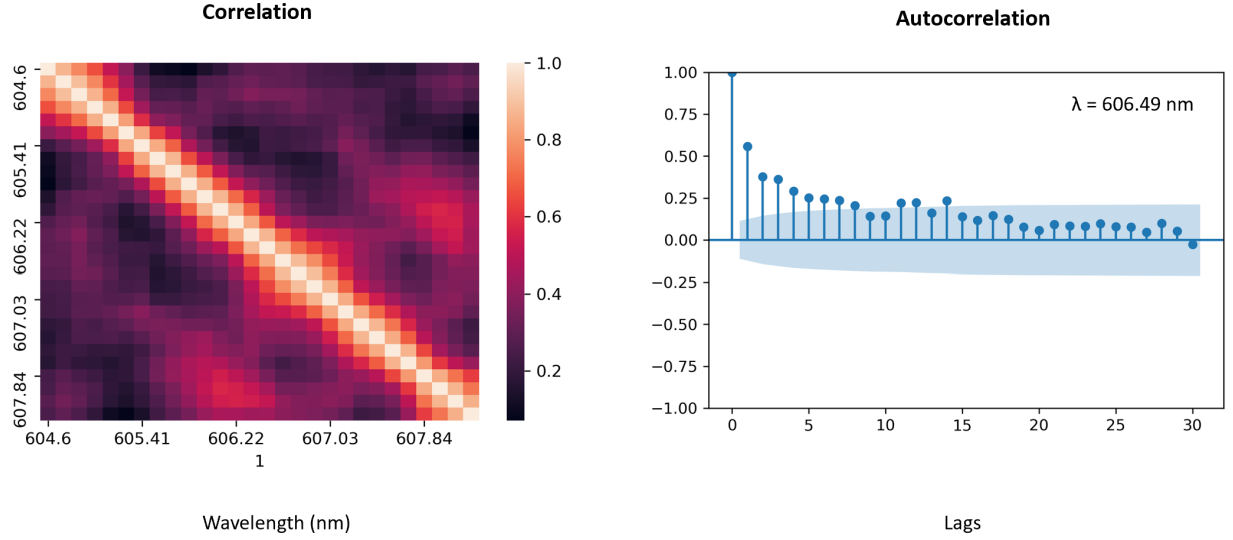


Figure 4.2: Correlation matrix and autocorrelation plot for band 1 and its center for 30 lags. Image on the left represents the correlation matrix for wavelengths withing band 1 and the image on the right represents the autocorrelation $\lambda = 606.49 \text{ nm}$.

Prediction

After checking the non-randomness nature of the data and developing the prediction models, I applied the trained models to the test sets and measured the RMSE. The measurement results are reported in Table 4.1. For each of the bands, the RMSE is calculated twice for LSTM prediction; once considering the model trained on full image (all the wavelengths) and another time considering the model trained only on the subset of wavelengths within the band. For almost all the experiments, the LSTM model does a better job than the shifted forecast; significantly for band 1 and band 2. The RMSE values measured for LSTM forecast show that higher correlations help the models to learn the historical data better, and therefore they lead to improvements in prediction results. To provide a more

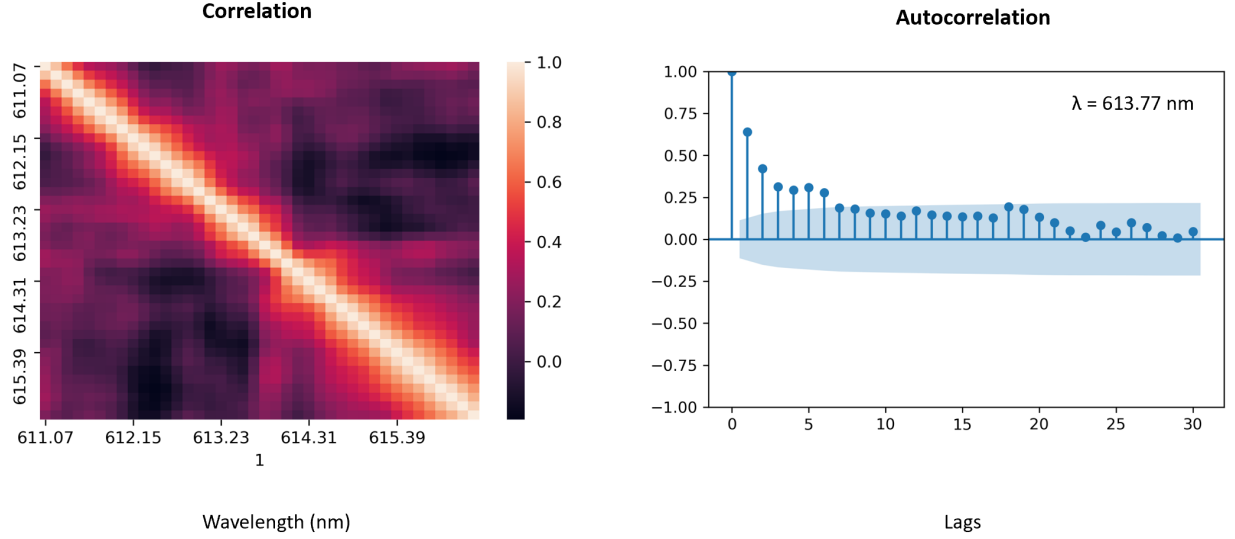


Figure 4.3: Correlation matrix and autocorrelation plot for band 2 and its center for 30 lags. Image on the left represents the correlation matrix for wavelengths withing band 1 and the image on the right represents the autocorrelation lambda=613.77 nm.

comprehensive evaluation of the performance of the predictive models, I measured the root mean square percentage error (RMSPE). While RMSE measures the average magnitude of the errors between predicted and actual values, RMSPE takes into account the relative percentage errors. These measurements are shown in Table 3.1. These measurements indicate consistent results when compared against the raw RMSE values reported for all the bands. However, when considering the full image, the RMSPE values differ. This disparity arises because the measurements for RMSPE take into account only non-zero pixels in the photos. The full image contains many zero values, whereas the bands contain mostly non-zero values. The RMSPE values show that the prediction error when using the LSTM method ranges from 6.5% to 24.66% for Bands 1-3. For Band 4, the simple shift forecast reports a smaller error.

In Fig. 4.6, Fig. 4.7, Fig. 4.8, Fig. 4.9, Fig. 4.10, I have displayed the prediction images using both LSTM forecast and the shifted forecast. For comparison of both models,

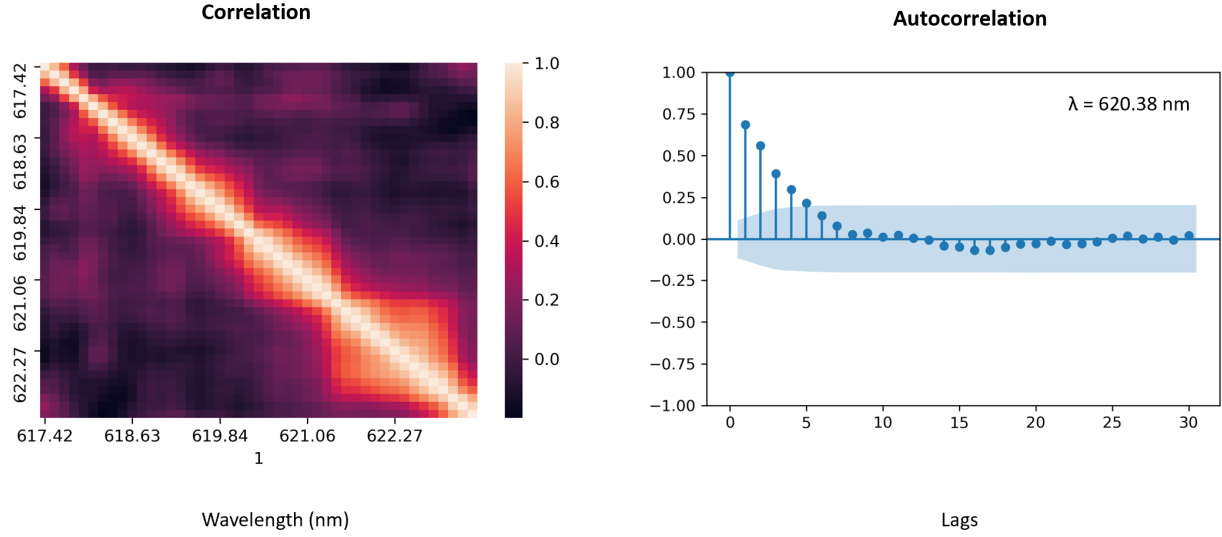


Figure 4.4: Correlation matrix and autocorrelation plot for band 3 and its center for 30 lags. Image on the left represents the correlation matrix for wavelengths withing band 1 and the image on the right represents the autocorrelation $\lambda=620.38$ nm.

I have shown the differences between the actual image and predicted image.

Fig. 4.11 shows the forecast results for two wavelengths belonging to band 3. For wavelength 620.52 nm, the LSTM model was able to fairly predict and follow the actual trend. For example it was able to predict several peaks and troughs, at the times 9-12s and 17-18s. For the wavelength of 619.17 nm, the LSTM model was able to detect the highest peak even though the predicted value is far from the actual data in terms of intensity. In both plots, there are many peaks and troughs that were not predicted by the LSTM model. Also note that the shifted forecast, by definition, identifies all peaks and troughs, but at a delayed time. The delay in QE prediction makes a simple shifted forecast infeasible to incorporate into a feedback control system.

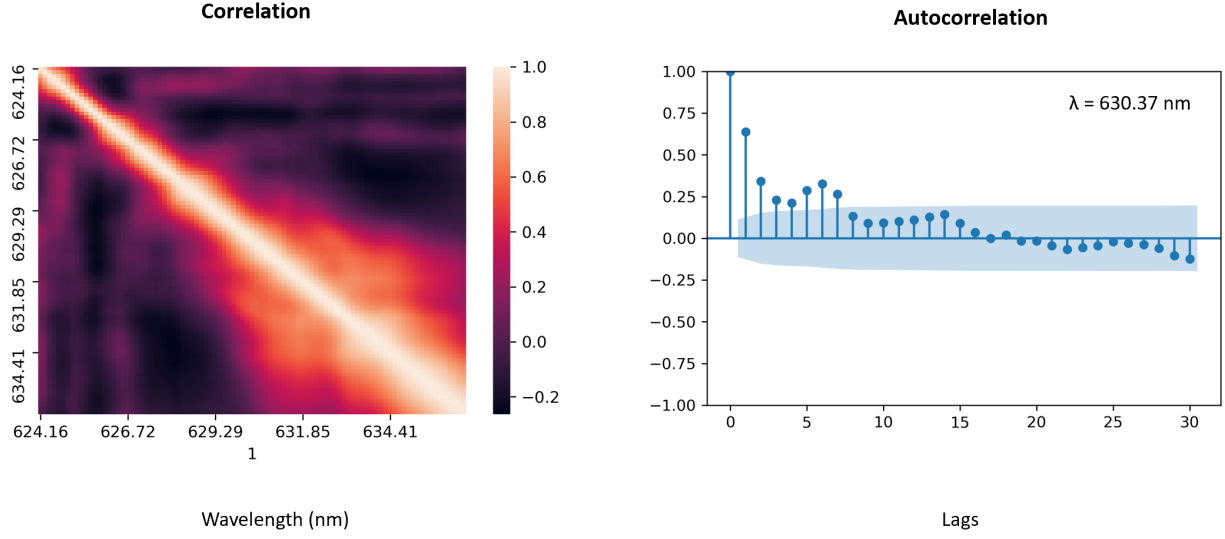


Figure 4.5: Correlation matrix and autocorrelation plot for band 4 and its center for 30 lags. Image on the left represents the correlation matrix for wavelengths withing band 1 and the image on the right represents the autocorrelation lambda=630.37 nm.

Table 4.1: Root Mean Square Error using LSTM and Shifted Forecast

	Wavelength Range (nm)	Number of Bins	Root Mean Square Error		
			LSTM (Band)	LSTM (Full Image)	Shift
Full Image	591.12 - 728.15	1023	79.12	79.12	79.40
Band 1	604.06 - 608.24	28	76.26	103.42	225.74
Band 2	611.07 - 616.34	40	313.60	336.80	387.19
Band 3	617.42 - 623.35	45	363.44	381.88	368.40
Band 4	624.16 - 636.57	93	305.54	326.70	296.05

Table 4.2: Root Mean Square Error Percentage using LSTM and Shifted Forecast

	Wavelength Range (nm)	Number of Bins	Root Mean Squared Percentage Error (in %)		
			LSTM (Band)	LSTM (Full Image)	Shift
Full Image	591.12 - 728.15	1023	372.62	372.62	71.76
Band 1	604.06 - 608.24	28	6.51	9.16	19.14
Band 2	611.07 - 616.34	40	19.58	21.53	24.80
Band 3	617.42 - 623.35	45	24.66	28.44	28.43
Band 4	624.16 - 636.57	93	37.81	39.87	29.71

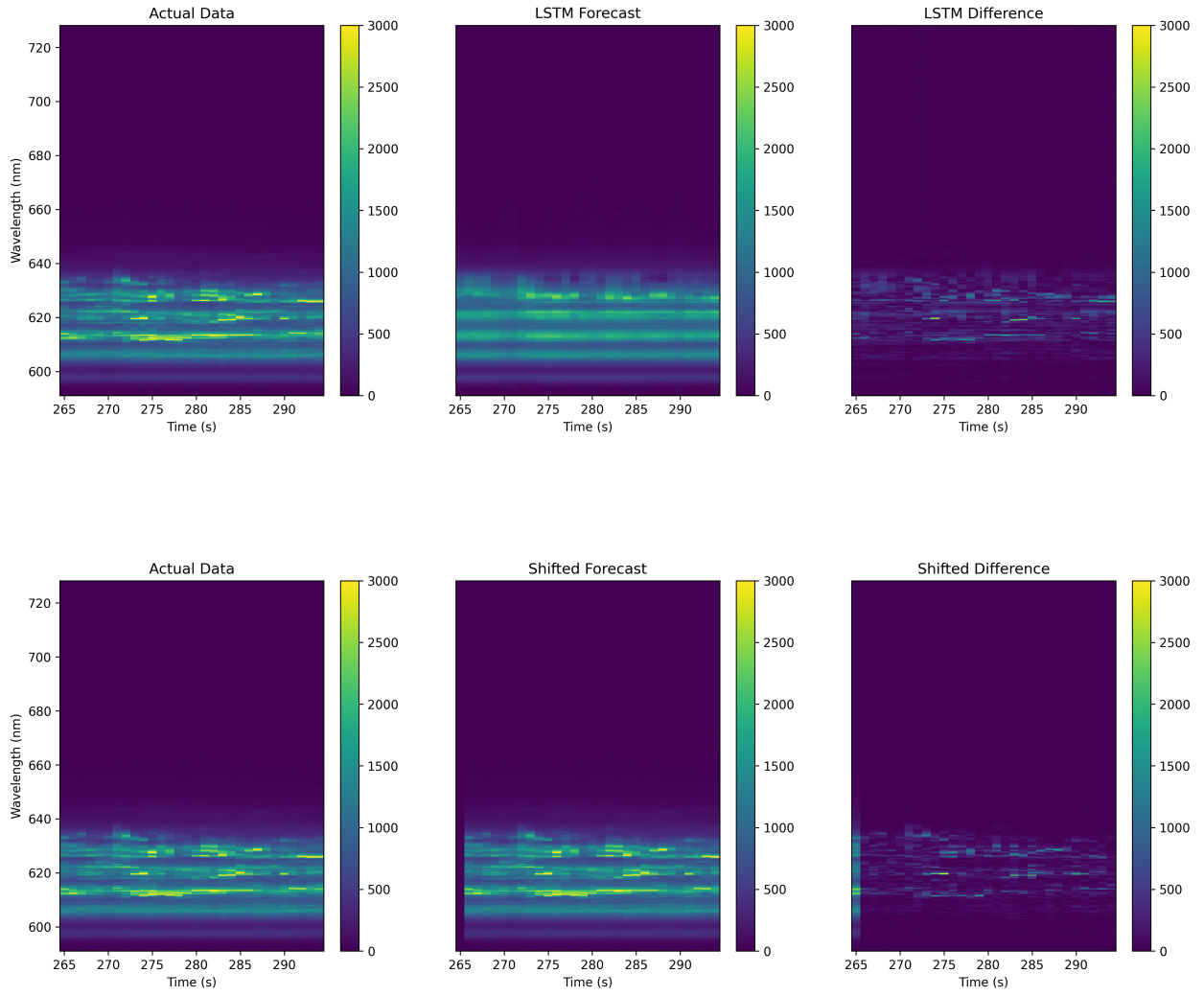


Figure 4.6: Prediction results using LSTM and shifted forecasts for the full image where all the wavelengths within the dataset are considered. Columns represent the actual image, forecast method, and difference between the actual image and the forecast. The first row represents the LSTM model and the second row represents the shifted forecast.

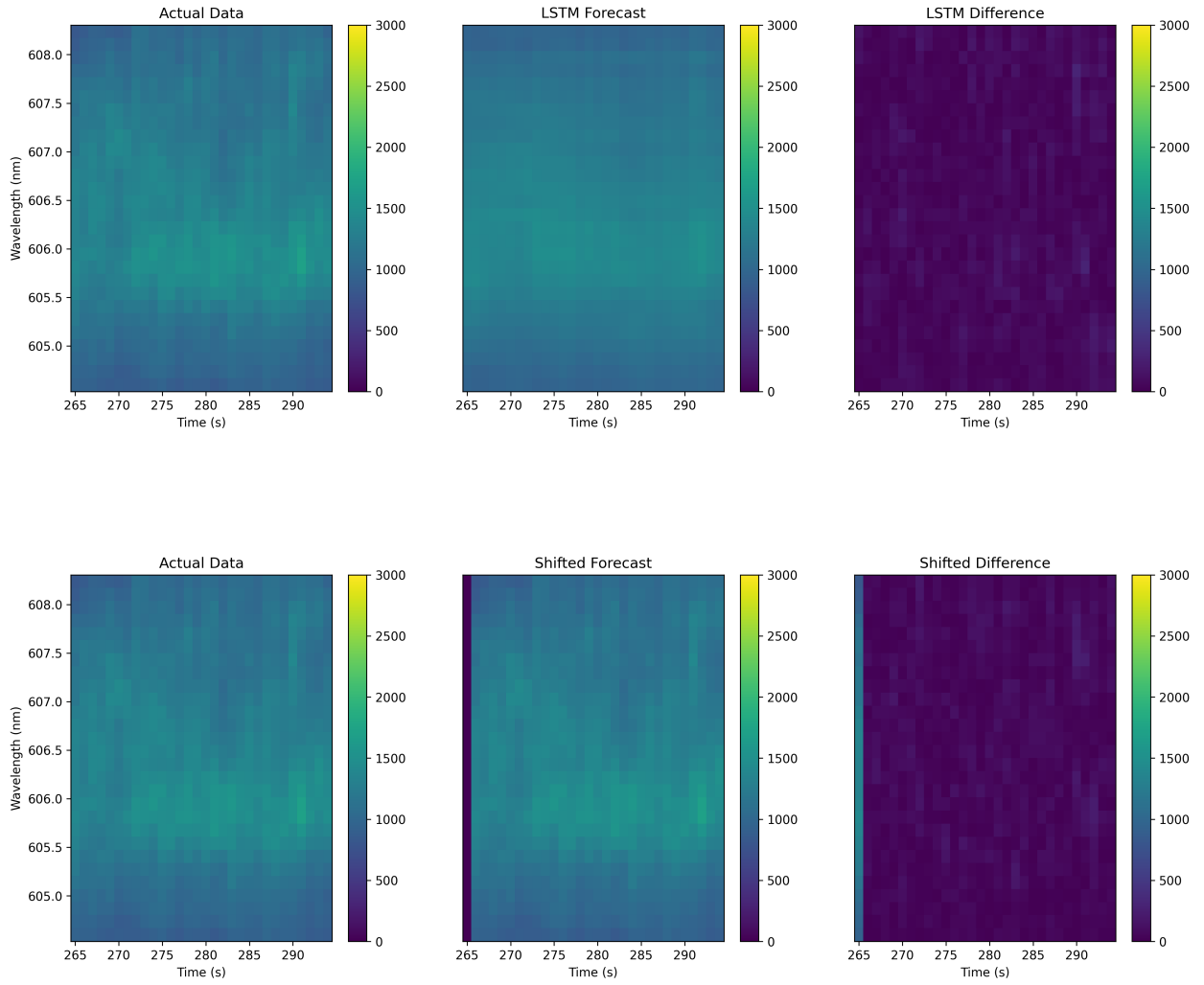


Figure 4.7: Prediction results using LSTM and shifted forecasts for band 1. Columns represent the actual image, forecast method, and difference between the actual image and the forecast. The first row represents the LSTM model and the second row represents the shifted forecast.

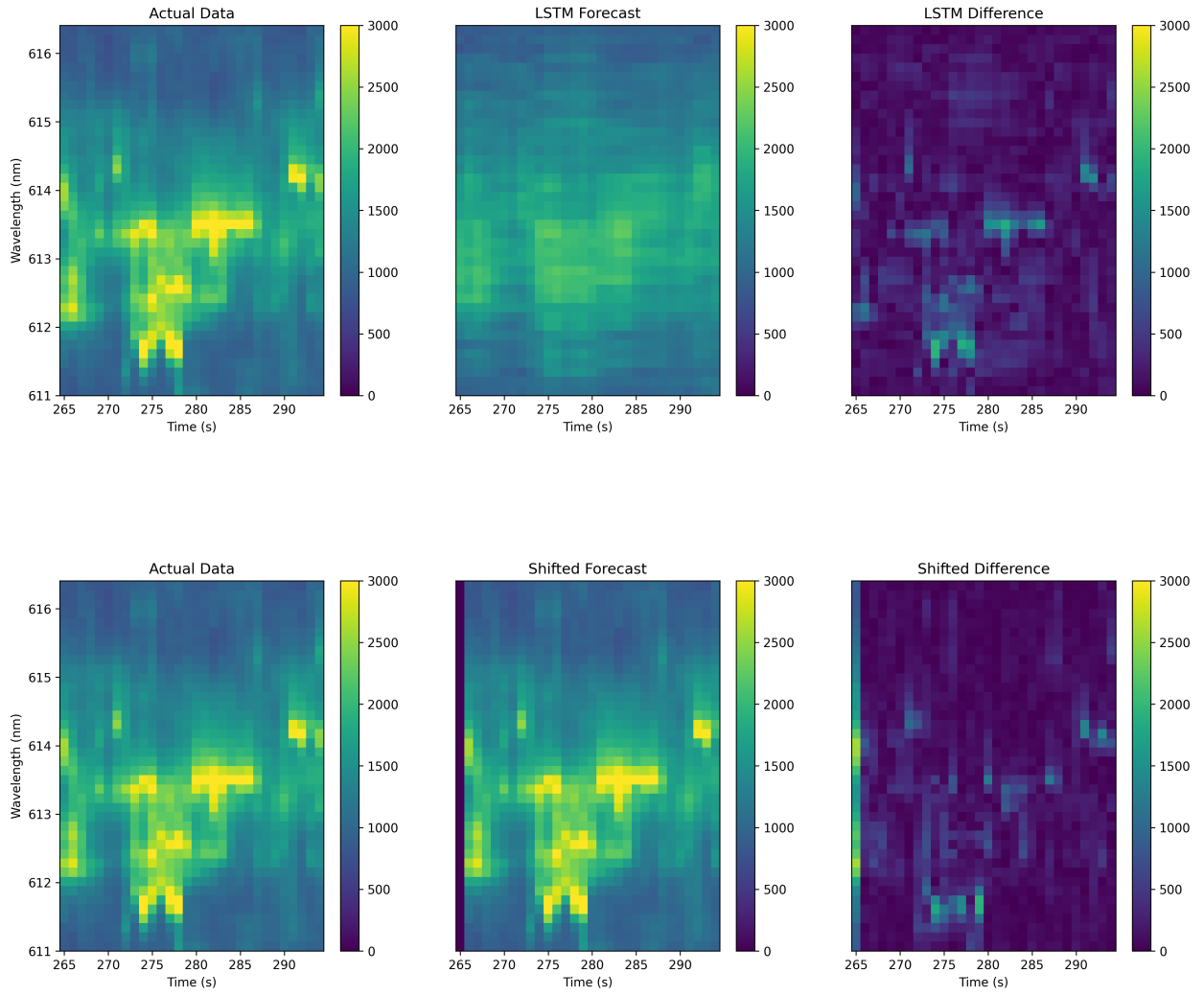


Figure 4.8: Prediction results using LSTM and shifted forecasts for band 2. Columns represent the actual image, forecast method, and difference between the actual image and the forecast. The first row represents the LSTM model and the second row represents the shifted forecast.

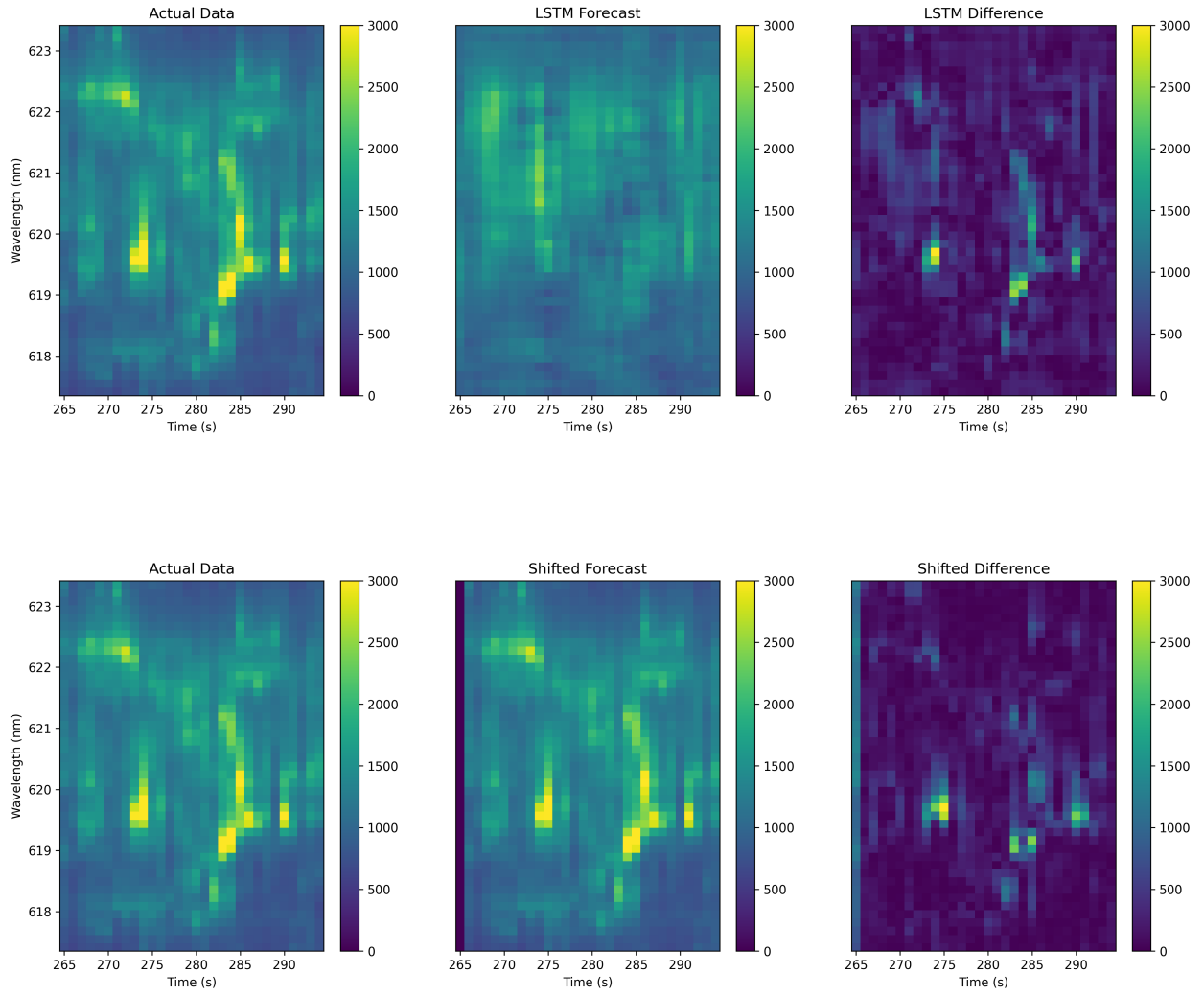


Figure 4.9: Prediction results using LSTM and shifted forecasts for band 3. Columns represent the actual image, forecast method, and difference between the actual image and the forecast. The first row represents the LSTM model and the second row represents the shifted forecast.

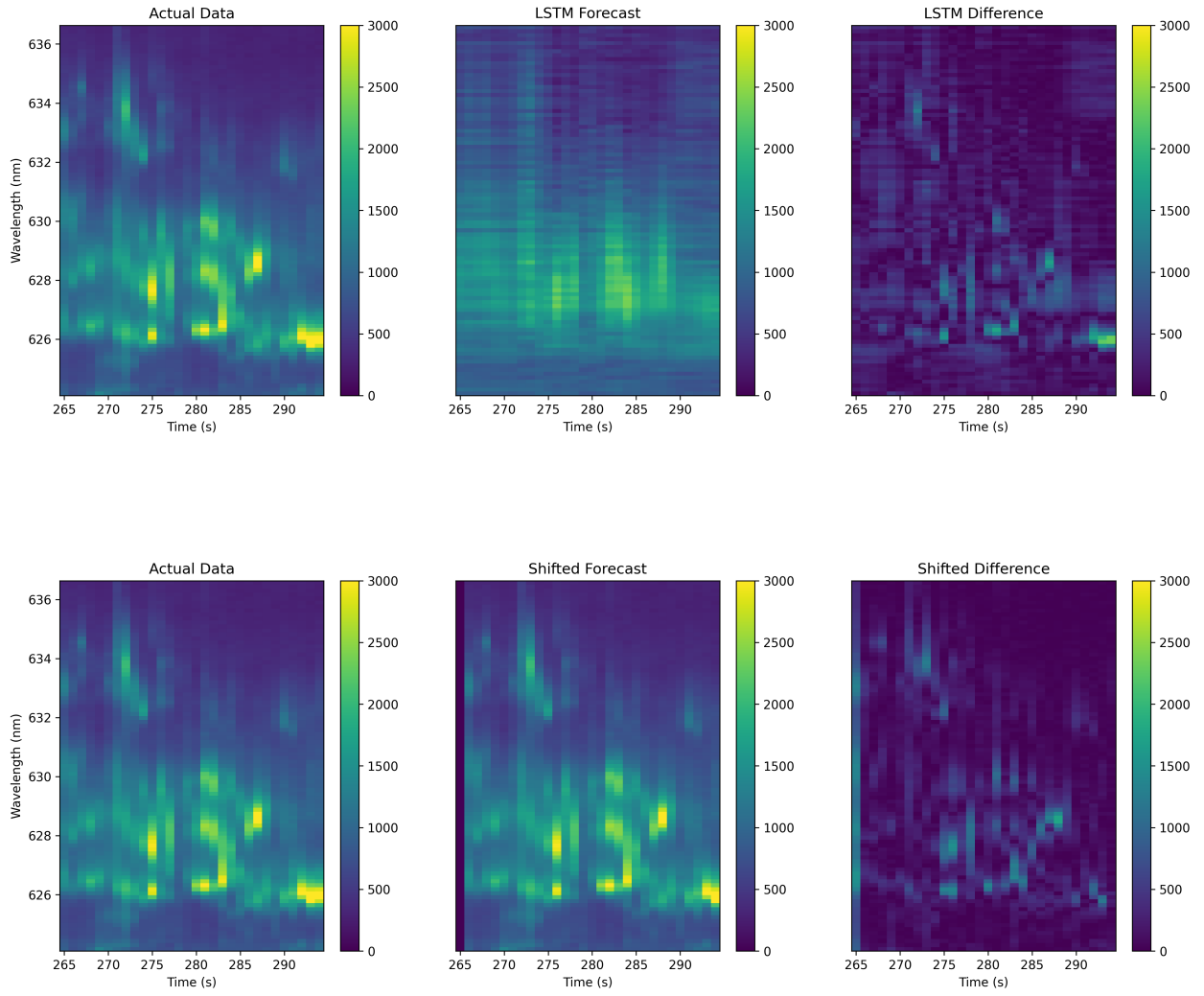


Figure 4.10: Prediction results using LSTM and shifted forecasts for band 4. Columns represent the actual image, forecast method, and difference between the actual image and the forecast. The first row represents the LSTM model and the second row represents the shifted forecast.

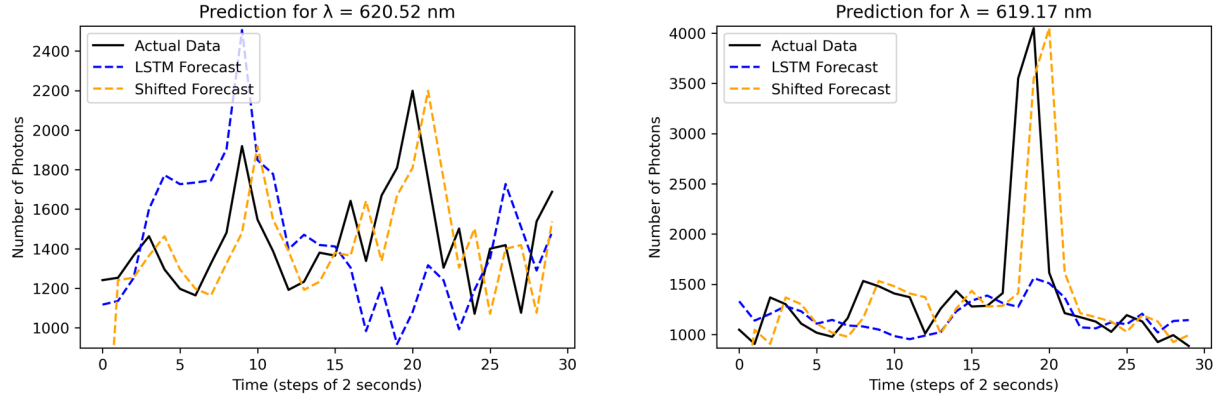


Figure 4.11: Single step prediction for wavelength 620.52 nm and 619.17 nm using LSTM forecast and Shifted forecast. For 620.52 nm, LSTM and shifted forecast RMSE are 420.37 and 346.7. For 619.17 nm, LSTM and shifted forecast RMSE are 648.80 and 654.57 respectively.

Discussion

To the best of my of knowledge, the development of high-performance models that can forecast the fluctuations of quantum emitters (and other nanoscale light sources) has not been studied prior to this work. Being able to forecast these fluctuations would provide deep insight into their origins and enable scientists and engineers to pursue predictive stabilization of nanoscale solid-state light sources. In my work, I developed LSTM-based models towards the prediction of QE fluctuations over time for emitters in 1L-WS₂ on gold. These emitters show dramatic fluctuations and providing an extreme test of the ability for neural network models to identify trends in the fluctuations. Despite the challenge of the task, my trained model was able to predict the general trend of a quantum emitter better than a simple shifted forecast, as measured by improved RMSE. To some extent, the model was able to predict the peaks and troughs of the fluctuations, though this behavior was inconsistent. I also showed that the fluctuations are of non-random nature, as measured by correlation

and autocorrelation. Building on this initial step, future work can focus on increasing the size and durations of the datasets on which the neural network is trained and explicitly training the neural network on wavelength and intensity fluctuations (as opposed to just the wavelength bands used here). These next steps will also lead to developing multi-step, multivariate time-series forecasting which could predict multiple steps into the future.

CONCLUSION

The fundamental properties of 2D materials have made them attractive for various device applications, including quantum computing. Different Van der Waals materials, such as hexagonal boron nitride (hBN), have been used to demonstrate alluring device applications, including emitting diodes, lasers, and optical modulators. The wavelength, polarization, and intensity of the photons from a quantum emission (QE) must be stable in order to optimize the utility and value of integrated quantum photonics. However, a significant issue for all solid-state single photon emitters is the random variation in emission energy brought on by the inhomogeneity of the immediate environment. In addition, small 2D flakes must be produced by an exfoliation technique and then discovered within microscope pictures in order to utilize these materials for study and product development. Manually carrying out these searches is a laborious and time-consuming process. This thesis has addressed critical challenges in identifying the 2D materials in microscope images, and assessed the random nature of the quantum fluctuations, and presented time-series forecasting deep learning models to analyze and predict QE fluctuations for the first time.

To tackle the issue of 2D flake detection, I have successfully implemented a novel deep learning pipeline to identify 2D materials in microscopic images. My analysis has demonstrated that this proposed pipeline is robust against variations in color and substrate background, and could be generalized to various microscope settings. One significant achievement of my research is the integration of my proposed method into the 2D quantum material pipeline (2D-QMaP), that has been under development by the MonArk Quantum Foundry. By doing this integration, I have provided the 2D-QMaP with fully automatic real-time 2D flake detection capabilities, a significant advancement in the field. The automated identification of 2D flakes, which was previously a laborious and time-consuming task, now enables researchers to streamline sample preparation and device fabrication processes for 2D

quantum materials research.

In addition to my contributions in automated 2D material identification, this thesis has ventured into the realm of quantum fluctuations analysis. I have explored the random nature of quantum fluctuations and introduced time-series forecasting deep learning models to analyze and predict QE fluctuations. My trained models can roughly track the actual trend of the data and, under certain data processing conditions, can forecast peaks and dips of the fluctuations. To the best of my knowledge, the development of high-performance models that can forecast the fluctuations of quantum emitters has not been studied prior to my work. The ability to predict these fluctuations can offer profound insights into their underlying causes, empowering scientists and engineers to work towards anticipatory measures for stabilizing nanoscale solid-state light sources.

In summary, my research has advanced the field of 2D quantum materials and paved the way for more efficient and accurate research and development processes. Through the automation of 2D material identification and the exploration of quantum fluctuations analysis, I have provided significant resources and perspectives that offer potentials for the creation of dependable quantum emission devices and the progress of quantum computing. This ultimately expands the horizons of what can be achieved in the realm of quantum technologies.

REFERENCES CITED

- [1] P. Ajayan, P. Kim, and K. Banerjee, “Two-dimensional van der waals materials,” *Physics Today*, vol. 69, pp. 38–44, 09 2016.
- [2] K. Novoselov and A. C. Neto, “Two-dimensional crystals-based heterostructures: materials with tailored properties,” *Physica Scripta*, vol. 2012, no. T146, p. 014006, 2012.
- [3] S. Masubuchi, E. Watanabe, Y. Seo, S. Okazaki, T. Sasagawa, K. Watanabe, T. Taniguchi, and T. Machida, “Deep-learning-based image segmentation integrated with optical microscopy for automatically searching for two-dimensional materials,” *npj 2D Materials and Applications*, vol. 4, no. 1, pp. 1–9, 2020.
- [4] R. Girshick, J. Donahue, T. Darrell, and J. Malik, “Rich feature hierarchies for accurate object detection and semantic segmentation,” in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2014, pp. 580–587.
- [5] R. Girshick, “Fast r-cnn,” in *Proceedings of the IEEE international conference on computer vision*, 2015, pp. 1440–1448.
- [6] K. He, G. Gkioxari, P. Dollár, and R. Girshick, “Mask r-cnn,” in *Proceedings of the IEEE international conference on computer vision*, 2017, pp. 2961–2969.
- [7] J. Redmon, S. Divvala, R. Girshick, and A. Farhadi, “You only look once: Unified, real-time object detection,” in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2016, pp. 779–788.
- [8] T.-Y. Lin, P. Goyal, R. Girshick, K. He, and P. Dollár, “Focal loss for dense object detection,” in *Proceedings of the IEEE international conference on computer vision*, 2017, pp. 2980–2988.
- [9] F. Sherratt, A. Plummer, and P. Iravani, “Understanding lstm network behaviour of imu-based locomotion mode recognition for applications in prostheses and wearables,” *Sensors*, vol. 21, no. 4, p. 1264, 2021.
- [10] S. Qiao, L.-C. Chen, and A. Yuille, “DetectoRS: Detecting objects with recursive feature pyramid and switchable atrous convolution,” in *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, 2021, pp. 10 213–10 224.
- [11] Q. H. Wang, K. Kalantar-Zadeh, A. Kis, J. N. Coleman, and M. S. Strano, “Electronics and optoelectronics of two-dimensional transition metal dichalcogenides,” *Nature nanotechnology*, vol. 7, no. 11, pp. 699–712, 2012.
- [12] X. Xu, W. Yao, D. Xiao, and T. F. Heinz, “Spin and pseudospins in layered transition metal dichalcogenides,” *Nature Physics*, vol. 10, no. 5, pp. 343–350, 2014.

- [13] G.-H. Lee, Y.-J. Yu, X. Cui, N. Petrone, C.-H. Lee, M. S. Choi, D.-Y. Lee, C. Lee, W. J. Yoo, K. Watanabe *et al.*, “Flexible and transparent mos2 field-effect transistors on hexagonal boron nitride-graphene heterostructures,” *ACS nano*, vol. 7, no. 9, pp. 7931–7936, 2013.
- [14] H. Shi, R. Yan, S. Bertolazzi, J. Brivio, B. Gao, A. Kis, D. Jena, H. G. Xing, and L. Huang, “Exciton dynamics in suspended monolayer and few-layer mos2 2d crystals,” *ACS nano*, vol. 7, no. 2, pp. 1072–1080, 2013.
- [15] R. Lv, J. A. Robinson, R. E. Schaak, D. Sun, Y. Sun, T. E. Mallouk, and M. Terrones, “Transition metal dichalcogenides and beyond: synthesis, properties, and applications of single-and few-layer nanosheets,” *Accounts of chemical research*, vol. 48, no. 1, pp. 56–64, 2015.
- [16] K. F. Mak and J. Shan, “Photonics and optoelectronics of 2d semiconductor transition metal dichalcogenides,” *Nature Photonics*, vol. 10, no. 4, pp. 216–226, 2016.
- [17] K. F. Mak, C. Lee, J. Hone, J. Shan, and T. F. Heinz, “Atomically thin mos 2: a new direct-gap semiconductor,” *Physical review letters*, vol. 105, no. 13, p. 136805, 2010.
- [18] A. Splendiani, L. Sun, Y. Zhang, T. Li, J. Kim, C.-Y. Chim, G. Galli, and F. Wang, “Emerging photoluminescence in monolayer mos2,” *Nano letters*, vol. 10, no. 4, pp. 1271–1275, 2010.
- [19] T. P. Darlington, C. Carmesin, M. Florian, E. Yanev, O. Ajayi, J. Ardelean, D. A. Rhodes, A. Ghiotto, A. Krayev, K. Watanabe *et al.*, “Imaging strain-localized excitons in nanoscale bubbles of monolayer wse2 at room temperature,” *Nature Nanotechnology*, vol. 15, no. 10, pp. 854–860, 2020.
- [20] Y. Li, B. Yang, S. Xu, B. Huang, and W. Duan, “Emergent phenomena in magnetic two-dimensional materials and van der waals heterostructures,” *ACS Applied Electronic Materials*, vol. 4, no. 7, pp. 3278–3302, 2022.
- [21] A. Gupta, T. Sakthivel, and S. Seal, “Recent development in 2d materials beyond graphene,” *Progress in Materials Science*, vol. 73, pp. 44–126, 2015.
- [22] K. Novoselov, o. A. Mishchenko, o. A. Carvalho, and A. Castro Neto, “2d materials and van der waals heterostructures,” *Science*, vol. 353, no. 6298, p. aac9439, 2016.
- [23] A. Lipatov, P. Chaudhary, Z. Guan, H. Lu, G. Li, O. Crégut, K. D. Dorkenoo, R. Proksch, S. Cherifi-Hertel, D.-F. Shao *et al.*, “Direct observation of ferroelectricity in two-dimensional mos2,” *npj 2D Materials and Applications*, vol. 6, no. 1, pp. 1–9, 2022.
- [24] X. Zhang, Y. Li, W. Mu, W. Bai, X. Sun, M. Zhao, Z. Zhang, F. Shan, and Z. Yang, “Advanced tape-exfoliated method for preparing large-area 2d monolayers: a review,” *2D Materials*, vol. 8, no. 3, p. 032002, 2021.

- [25] H.-W. Guo, Z. Hu, Z.-B. Liu, and J.-G. Tian, “Stacking of 2d materials,” *Advanced Functional Materials*, vol. 31, no. 4, p. 2007810, 2021.
- [26] Z. Zou, Z. Shi, Y. Guo, and J. Ye, “Object detection in 20 years: A survey,” *arXiv preprint arXiv:1905.05055*, 2019.
- [27] L. Liu, W. Ouyang, X. Wang, P. Fieguth, J. Chen, X. Liu, and M. Pietikäinen, “Deep learning for generic object detection: A survey,” *International journal of computer vision*, vol. 128, no. 2, pp. 261–318, 2020.
- [28] A. R. Pathak, M. Pandey, and S. Rautaray, “Application of deep learning for object detection,” *Procedia computer science*, vol. 132, pp. 1706–1717, 2018.
- [29] N. O’Mahony, S. Campbell, A. Carvalho, S. Harapanahalli, G. V. Hernandez, L. Krpalkova, D. Riordan, and J. Walsh, “Deep learning vs. traditional computer vision,” in *Science and information conference*. Springer, 2019, pp. 128–144.
- [30] B. Han, Y. Lin, Y. Yang, N. Mao, W. Li, H. Wang, K. Yasuda, X. Wang, V. Fatemi, L. Zhou *et al.*, “Deep-learning-enabled fast optical identification and characterization of 2d materials,” *Advanced Materials*, vol. 32, no. 29, p. 2000953, 2020.
- [31] X. Lin, Z. Si, W. Fu, J. Yang, S. Guo, Y. Cao, J. Zhang, X. Wang, P. Liu, K. Jiang *et al.*, “Intelligent identification of two-dimensional nanostructures by machine-learning optical microscopy,” *Nano Research*, vol. 11, no. 12, pp. 6316–6324, 2018.
- [32] J. Yang and H. Yao, “Automated identification and characterization of two-dimensional materials via machine learning-based processing of optical microscope images,” *Extreme Mechanics Letters*, vol. 39, p. 100771, 2020.
- [33] Y. Li, Y. Kong, J. Peng, C. Yu, Z. Li, P. Li, Y. Liu, C.-F. Gao, and R. Wu, “Rapid identification of two-dimensional materials via machine learning assisted optical microscopy,” *Journal of Materiomics*, vol. 5, no. 3, pp. 413–421, 2019.
- [34] S. Masubuchi and T. Machida, “Classifying optical microscope images of exfoliated graphene flakes by data-driven machine learning,” *npj 2D Materials and Applications*, vol. 3, no. 1, pp. 1–7, 2019.
- [35] X. B. Nguyen, A. Bisht, H. Churchill, and K. Luu, “Two-dimensional quantum material identification via self-attention and soft-labeling in deep learning,” *arXiv preprint arXiv:2205.15948*, 2022.
- [36] M. Mudelsee, “Trend analysis of climate time series: A review of methods,” *Earth-science reviews*, vol. 190, pp. 310–322, 2019.
- [37] D. S. Stoffer and H. Ombao, “Special issue on time series analysis in the biological sciences,” pp. 701–703, 2012.

- [38] E. J. Topol, “High-performance medicine: the convergence of human and artificial intelligence,” *Nature medicine*, vol. 25, no. 1, pp. 44–56, 2019.
- [39] J.-H. Böse, V. Flunkert, J. Gasthaus, T. Januschowski, D. Lange, D. Salinas, S. Schelter, M. Seeger, and Y. Wang, “Probabilistic demand forecasting at scale,” *Proceedings of the VLDB Endowment*, vol. 10, no. 12, pp. 1694–1705, 2017.
- [40] T. G. Andersen, T. Bollerslev, P. Christoffersen, and F. X. Diebold, “Volatility forecasting,” 2005.
- [41] B. Lim and S. Zohren, “Time-series forecasting with deep learning: a survey,” *Philosophical Transactions of the Royal Society A*, vol. 379, no. 2194, p. 20200209, 2021.
- [42] A. K. Geim and A. H. MacDonald, “Graphene: Exploring carbon flatland,” *Physics Today*, vol. 60, no. 8, pp. 35–41, 2007. [Online]. Available: <https://doi.org/10.1063/1.2774096>
- [43] K. Zhang, Y. Feng, F. Wang, Z. Yang, and J. Wang, “Two dimensional hexagonal boron nitride (2d-hbn): synthesis, properties and applications,” *Journal of Materials Chemistry C*, vol. 5, no. 46, pp. 11 992–12 022, 2017.
- [44] H. Jiang and J. Lin, “Hexagonal boron nitride for deep ultraviolet photonic devices,” *Semiconductor Science and Technology*, vol. 29, no. 8, p. 084003, 2014.
- [45] F. Hui, C. Pan, Y. Shi, Y. Ji, E. Grustan-Gutierrez, and M. Lanza, “On the use of two dimensional hexagonal boron nitride as dielectric,” *Microelectronic Engineering*, vol. 163, pp. 119–133, 2016.
- [46] G. C. Constantinescu and N. D. Hine, “Multipurpose black-phosphorus/hbn heterostructures,” *Nano letters*, vol. 16, no. 4, pp. 2586–2594, 2016.
- [47] J. Bao, M. Edwards, S. Huang, Y. Zhang, Y. Fu, X. Lu, Z. Yuan, K. Jeppson, and J. Liu, “Two-dimensional hexagonal boron nitride as lateral heat spreader in electrically insulating packaging,” *Journal of Physics D: Applied Physics*, vol. 49, no. 26, p. 265501, 2016.
- [48] D. Chimene, D. L. Alge, and A. K. Gaharwar, “Two-dimensional nanomaterials for biomedical applications: emerging trends and future prospects,” *Advanced Materials*, vol. 27, no. 45, pp. 7261–7284, 2015.
- [49] T. Knobloch, Y. Y. Illarionov, F. Ducry, C. Schleich, S. Wachter, T. Müller, M. Waltl, M. Lanza, M. I. Vexler, M. Luisier *et al.*, “On the suitability of hbn as an insulator for 2d material-based ultrascaled cmos devices,” *arXiv preprint arXiv:2008.04144*, 2020.

- [50] P. K. Kannan, D. J. Late, H. Morgan, and C. S. Rout, “Recent developments in 2d layered inorganic nanomaterials for sensing,” *Nanoscale*, vol. 7, no. 32, pp. 13 293–13 312, 2015.
- [51] S. Yu, X. Wu, Y. Wang, X. Guo, and L. Tong, “2d materials for optical modulation: challenges and opportunities,” *Advanced Materials*, vol. 29, no. 14, p. 1606128, 2017.
- [52] M. Osada and T. Sasaki, “Two-dimensional dielectric nanosheets: novel nanoelectronics from nanocrystal building blocks,” *Advanced Materials*, vol. 24, no. 2, pp. 210–228, 2012.
- [53] K. Kalantar-zadeh, J. Z. Ou, T. Daeneke, M. S. Strano, M. Pumera, and S. L. Gras, “Two-dimensional transition metal dichalcogenides in biosystems,” *Advanced Functional Materials*, vol. 25, no. 32, pp. 5086–5099, 2015.
- [54] A. C. Ferrari, F. Bonaccorso, V. Fal’Ko, K. S. Novoselov, S. Roche, P. Bøggild, S. Borini, F. H. Koppens, V. Palermo, N. Pugno *et al.*, “Science and technology roadmap for graphene, related two-dimensional crystals, and hybrid systems,” *Nanoscale*, vol. 7, no. 11, pp. 4598–4810, 2015.
- [55] H. Zhang, “Ultrathin two-dimensional nanomaterials,” *ACS nano*, vol. 9, no. 10, pp. 9451–9469, 2015.
- [56] K. Xu, P. Chen, X. Li, Y. Tong, H. Ding, X. Wu, W. Chu, Z. Peng, C. Wu, and Y. Xie, “Metallic nickel nitride nanosheets realizing enhanced electrochemical water oxidation,” *Journal of the American Chemical Society*, vol. 137, no. 12, pp. 4119–4125, 2015.
- [57] X. Sun, H. Deng, W. Zhu, Z. Yu, C. Wu, and Y. Xie, “Interface engineering in two-dimensional heterostructures: Towards an advanced catalyst for ullmann couplings,” *Angewandte Chemie International Edition*, vol. 55, no. 5, pp. 1704–1709, 2016.
- [58] M. Chhowalla, H. S. Shin, G. Eda, L.-J. Li, K. P. Loh, and H. Zhang, “The chemistry of two-dimensional layered transition metal dichalcogenide nanosheets,” *Nature chemistry*, vol. 5, no. 4, pp. 263–275, 2013.
- [59] X. Peng, L. Peng, C. Wu, and Y. Xie, “Two dimensional nanomaterials for flexible supercapacitors,” *Chemical Society Reviews*, vol. 43, no. 10, pp. 3303–3323, 2014.
- [60] W. Li, X. Qian, and J. Li, “Phase transitions in 2d materials,” *Nature Reviews Materials*, vol. 6, no. 9, pp. 829–846, 2021.
- [61] M. Gibertini, M. Koperski, A. F. Morpurgo, and K. S. Novoselov, “Magnetic 2d materials and heterostructures,” *Nature nanotechnology*, vol. 14, no. 5, pp. 408–419, 2019.

- [62] E. Gao, S.-Z. Lin, Z. Qin, M. J. Buehler, X.-Q. Feng, and Z. Xu, “Mechanical exfoliation of two-dimensional materials,” *Journal of the Mechanics and Physics of Solids*, vol. 115, pp. 248–262, 2018.
- [63] R. Hergett, “Quantum foundation,” May 2022. [Online]. Available: <https://www.montana.edu/news/mountainsandminds/22122/quantum-foundation>
- [64] N. Gisin, G. Ribordy, W. Tittel, and H. Zbinden, “Quantum cryptography,” *Reviews of modern physics*, vol. 74, no. 1, p. 145, 2002.
- [65] T. D. Ladd, F. Jelezko, R. Laflamme, Y. Nakamura, C. Monroe, and J. L. O’Brien, “Quantum computers,” *nature*, vol. 464, no. 7285, pp. 45–53, 2010.
- [66] A. Aspuru-Guzik and P. Walther, “Photonic quantum simulators,” *Nature physics*, vol. 8, no. 4, pp. 285–291, 2012.
- [67] I. Aharonovich and M. Toth, “Quantum emitters in two dimensions,” *Science*, vol. 358, no. 6360, pp. 170–171, 2017.
- [68] J. Wang, F. Sciarrino, A. Laing, and M. G. Thompson, “Integrated photonic quantum technologies,” *Nature Photonics*, vol. 14, no. 5, pp. 273–284, 2020.
- [69] A. K. Geim and I. V. Grigorieva, “Van der waals heterostructures,” *Nature*, vol. 499, no. 7459, pp. 419–425, 2013.
- [70] R. Sundaram, M. Engel, A. Lombardo, R. Krupke, A. Ferrari, P. Avouris, and M. Steiner, “Electroluminescence in single layer mos2,” *Nano letters*, vol. 13, no. 4, pp. 1416–1421, 2013.
- [71] Y. Ye, Z. J. Wong, X. Lu, X. Ni, H. Zhu, X. Chen, Y. Wang, and X. Zhang, “Monolayer excitonic laser,” *Nature Photonics*, vol. 9, no. 11, pp. 733–737, 2015.
- [72] M. Liu, X. Yin, E. Ulin-Avila, B. Geng, T. Zentgraf, L. Ju, F. Wang, and X. Zhang, “A graphene-based broadband optical modulator,” *Nature*, vol. 474, no. 7349, pp. 64–67, 2011.
- [73] Y. Xia, Q. Li, J. Kim, W. Bao, C. Gong, S. Yang, Y. Wang, and X. Zhang, “Room-temperature giant stark effect of single photon emitter in van der waals material,” *Nano letters*, vol. 19, no. 10, pp. 7100–7105, 2019.
- [74] M. R. Bachute and J. M. Subhedar, “Autonomous driving architectures: insights of machine learning and deep learning algorithms,” *Machine Learning with Applications*, vol. 6, p. 100164, 2021.
- [75] J. Wiens and E. S. Shenoy, “Machine learning for healthcare: on the verge of a major shift in healthcare epidemiology,” *Clinical Infectious Diseases*, vol. 66, no. 1, pp. 149–153, 2018.

- [76] K. Shaukat, S. Luo, V. Varadharajan, I. A. Hameed, and M. Xu, "A survey on machine learning techniques for cyber security in the last decade," *IEEE access*, vol. 8, pp. 222 310–222 354, 2020.
- [77] A. Rajkomar, J. Dean, and I. Kohane, "Machine learning in medicine," *New England Journal of Medicine*, vol. 380, no. 14, pp. 1347–1358, 2019.
- [78] P. K. Donepudi, "Machine learning and artificial intelligence in banking," *Engineering International*, vol. 5, no. 2, pp. 83–86, 2017.
- [79] A. Dogan and D. Birant, "Machine learning and data mining in manufacturing," *Expert Systems with Applications*, vol. 166, p. 114060, 2021.
- [80] J. E. Van Engelen and H. H. Hoos, "A survey on semi-supervised learning," *Machine learning*, vol. 109, no. 2, pp. 373–440, 2020.
- [81] P. Ladosz, L. Weng, M. Kim, and H. Oh, "Exploration in deep reinforcement learning: A survey," *Information Fusion*, 2022.
- [82] K. Marias, "The constantly evolving role of medical image processing in oncology: From traditional medical image processing to imaging biomarkers and radiomics," *Journal of imaging*, vol. 7, no. 8, p. 124, 2021.
- [83] R. B. Hegde, K. Prasad, H. Hebbar, and B. M. K. Singh, "Comparison of traditional image processing and deep learning approaches for classification of white blood cells in peripheral blood smear images," *Biocybernetics and Biomedical Engineering*, vol. 39, no. 2, pp. 382–392, 2019.
- [84] S. D. Jawak, P. Devliyal, A. J. Luis *et al.*, "A comprehensive review on pixel oriented and object oriented methods for information extraction from remotely sensed satellite images with a special emphasis on cryospheric applications," *Advances in Remote Sensing*, vol. 4, no. 03, p. 177, 2015.
- [85] H. Ramezani, F. Keynia, and F. Ramezani, "A novel image steganography in contourletdomain using genetic algorithm," *International Journal of Future Computer and Communication*, vol. 2, no. 4, p. 359, 2013.
- [86] M. P. Ekstrom, *Digital image processing techniques*. Academic Press, 2012, vol. 2.
- [87] M. S. Alkoffash, M. J. Bawaneh, H. Muaidi, S. Alqrainy, and M. Alzghool, "A survey of digital image processing techniques in character recognition," *International Journal of Computer Science and Network Security (IJCSNS)*, vol. 14, no. 3, p. 65, 2014.
- [88] S. Krig and S. Krig, "Image pre-processing," *Computer Vision Metrics: Textbook Edition*, pp. 35–74, 2016.

- [89] H. Ramezani, M. Mohammadi, and A. S. Molahoseini, "An efficient implementation of low-latency two-dimensional gaussian smoothing filter using approximate carry-save adder," *Journal of Circuits, Systems and Computers*, vol. 30, no. 02, p. 2150021, 2021.
- [90] N. Otsu, "A threshold selection method from gray-level histograms," *IEEE transactions on systems, man, and cybernetics*, vol. 9, no. 1, pp. 62–66, 1979.
- [91] M. L. Comer and E. J. Delp III, "Morphological operations for color image processing," *Journal of electronic imaging*, vol. 8, no. 3, pp. 279–289, 1999.
- [92] T. Blaschke, C. Burnett, and A. Pekkarinen, "Image segmentation methods for object-based analysis and classification," *Remote sensing image analysis: Including the spatial domain*, pp. 211–236, 2004.
- [93] P. Dhankhar and N. Sahu, "A review and research of edge detection techniques for image segmentation," *International Journal of Computer Science and Mobile Computing*, vol. 2, no. 7, pp. 86–92, 2013.
- [94] H. G. Kaganami and Z. Bei, "Region-based segmentation versus edge detection," in *2009 Fifth International Conference on Intelligent Information Hiding and Multimedia Signal Processing*. IEEE, 2009, pp. 1217–1221.
- [95] S. Karthick, K. Sathiyasekar, and A. Puraneeswari, "A survey based on region based segmentation," *International Journal of Engineering Trends and Technology*, vol. 7, no. 3, pp. 143–147, 2014.
- [96] A. Singh, N. Thakur, and A. Sharma, "A review of supervised machine learning algorithms," in *2016 3rd international conference on computing for sustainable global development (INDIACom)*. Ieee, 2016, pp. 1310–1315.
- [97] P. Viola and M. Jones, "Rapid object detection using a boosted cascade of simple features," *IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 2001.
- [98] P. Viola and M. J. Jones, "Robust real-time face detection," *International journal of computer vision*, vol. 57, pp. 137–154, 2004.
- [99] N. Dalal and B. Triggs, "Histograms of oriented gradients for human detection," in *2005 IEEE computer society conference on computer vision and pattern recognition (CVPR'05)*, vol. 1. Ieee, 2005, pp. 886–893.
- [100] P. Felzenszwalb, D. McAllester, and D. Ramanan, "A discriminatively trained, multiscale, deformable part model," in *2008 IEEE conference on computer vision and pattern recognition*. Ieee, 2008, pp. 1–8.

- [101] P. F. Felzenszwalb, R. B. Girshick, and D. McAllester, “Cascade object detection with deformable part models,” in *2010 IEEE Computer society conference on computer vision and pattern recognition*. Ieee, 2010, pp. 2241–2248.
- [102] T. Malisiewicz, A. Gupta, and A. A. Efros, “Ensemble of exemplar-svms for object detection and beyond,” in *2011 International conference on computer vision*. IEEE, 2011, pp. 89–96.
- [103] Z. Zou, K. Chen, Z. Shi, Y. Guo, and J. Ye, “Object detection in 20 years: A survey,” *Proceedings of the IEEE*, vol. 111, no. 3, pp. 257–276, 2023.
- [104] P. F. Felzenszwalb, R. B. Girshick, D. McAllester, and D. Ramanan, “Object detection with discriminatively trained part-based models,” *IEEE transactions on pattern analysis and machine intelligence*, vol. 32, no. 9, pp. 1627–1645, 2009.
- [105] R. Girshick, P. Felzenszwalb, and D. McAllester, “Object detection with grammar models,” *Advances in neural information processing systems*, vol. 24, 2011.
- [106] R. B. Girshick, *From rigid templates to grammars: Object detection with structured models*. The University of Chicago, 2012.
- [107] X. Wu, D. Sahoo, and S. C. Hoi, “Recent advances in deep learning for object detection,” *Neurocomputing*, vol. 396, pp. 39–64, 2020.
- [108] S. Ren, K. He, R. Girshick, and J. Sun, “Faster r-cnn: Towards real-time object detection with region proposal networks,” *Advances in neural information processing systems*, vol. 28, 2015.
- [109] J. Redmon and A. Farhadi, “Yolo9000: better, faster, stronger,” in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2017, pp. 7263–7271.
- [110] —, “Yolov3: An incremental improvement,” *arXiv preprint arXiv:1804.02767*, 2018.
- [111] A. Bochkovskiy, C.-Y. Wang, and H.-Y. M. Liao, “Yolov4: Optimal speed and accuracy of object detection,” *arXiv preprint arXiv:2004.10934*, 2020.
- [112] W. Liu, D. Anguelov, D. Erhan, C. Szegedy, S. Reed, C.-Y. Fu, and A. C. Berg, “Ssd: Single shot multibox detector,” in *European conference on computer vision*. Springer, 2016, pp. 21–37.
- [113] T.-Y. Lin, P. Dollár, R. Girshick, K. He, B. Hariharan, and S. Belongie, “Feature pyramid networks for object detection,” in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2017, pp. 2117–2125.
- [114] K. He, X. Zhang, S. Ren, and J. Sun, “Deep residual learning for image recognition,” in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2016, pp. 770–778.

- [115] K. P. Sinaga and M.-S. Yang, “Unsupervised k-means clustering algorithm,” *IEEE access*, vol. 8, pp. 80 716–80 727, 2020.
- [116] G. J. McLachlan and K. E. Basford, *Mixture models: Inference and applications to clustering*. M. Dekker New York, 1988, vol. 38.
- [117] L. Kaufman and P. J. Rousseeuw, *Finding groups in data: an introduction to cluster analysis*. John Wiley & Sons, 2009.
- [118] A. K. Jain, “Data clustering: 50 years beyond k-means,” *Pattern recognition letters*, vol. 31, no. 8, pp. 651–666, 2010.
- [119] M. Köppen, “The curse of dimensionality,” in *5th online world conference on soft computing in industrial applications (WSC5)*, vol. 1, 2000, pp. 4–8.
- [120] S. Wold, K. Esbensen, and P. Geladi, “Principal component analysis,” *Chemometrics and intelligent laboratory systems*, vol. 2, no. 1-3, pp. 37–52, 1987.
- [121] F. Ramezani, C. Major, C. Barney, J. Williams, B. J. LaMeres, and B. M. Whitaker, “Identifying patterns in fault recovery techniques and hardware status of radiation tolerant computers using principal components analysis,” in *2022 Intermountain Engineering, Technology and Computing (IETC)*. IEEE, 2022, pp. 1–6.
- [122] K. Liyanage, F. Ramezani, and B. M. Whitaker, “Covid-19 chest ct scan image classification using lcksvd and frozen sparse coding,” in *Proceedings of 2021 International Conference on Medical Imaging and Computer-Aided Diagnosis (MICAD 2021) Medical Imaging and Computer-Aided Diagnosis*. Springer, 2022, pp. 272–281.
- [123] I. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville, and Y. Bengio, “Generative adversarial nets,” in *Advances in Neural Information Processing Systems*, Z. Ghahramani, M. Welling, C. Cortes, N. Lawrence, and K. Weinberger, Eds., vol. 27. Curran Associates, Inc., 2014. [Online]. Available: https://proceedings.neurips.cc/paper_files/paper/2014/file/5ca3e9b122f61f8f06494c97b1afccf3-Paper.pdf
- [124] A. Creswell, T. White, V. Dumoulin, K. Arulkumaran, B. Sengupta, and A. A. Bharath, “Generative adversarial networks: An overview,” *IEEE signal processing magazine*, vol. 35, no. 1, pp. 53–65, 2018.
- [125] K. Wang, C. Gou, Y. Duan, Y. Lin, X. Zheng, and F.-Y. Wang, “Generative adversarial networks: introduction and outlook,” *IEEE/CAA Journal of Automatica Sinica*, vol. 4, no. 4, pp. 588–598, 2017.
- [126] M. Carranza-García, J. Torres-Mateo, P. Lara-Benítez, and J. García-Gutiérrez, “On the performance of one-stage and two-stage object detectors in autonomous vehicles using camera data,” *Remote Sensing*, vol. 13, no. 1, p. 89, 2020.

- [127] R. Simhambhatla, K. Okiah, S. Kuchkula, and R. Slater, "Self-driving cars: Evaluation of deep learning techniques for object detection in different driving conditions," *SMU Data Science Review*, vol. 2, no. 1, p. 23, 2019.
- [128] A. Gupta, A. Anpalagan, L. Guan, and A. S. Khwaja, "Deep learning for object detection and scene perception in self-driving cars: Survey, challenges, and open issues," *Array*, vol. 10, p. 100057, 2021.
- [129] R. Kulkarni, S. Dhavalikar, and S. Bangar, "Traffic light detection and recognition for self driving cars using deep learning," in *2018 Fourth International Conference on Computing Communication Control and Automation (ICCCUBEA)*. IEEE, 2018, pp. 1–4.
- [130] A. Kumar, A. Kaur, and M. Kumar, "Face detection techniques: a review," *Artificial Intelligence Review*, vol. 52, no. 2, pp. 927–948, 2019.
- [131] Z. Ren, S. Yang, F. Zou, F. Yang, C. Luan, and K. Li, "A face tracking framework based on convolutional neural networks and kalman filter," in *2017 8th IEEE international conference on software engineering and service science (ICSESS)*. IEEE, 2017, pp. 410–413.
- [132] H. Jiang and E. Learned-Miller, "Face detection with the faster r-cnn," in *2017 12th IEEE international conference on automatic face & gesture recognition (FG 2017)*. IEEE, 2017, pp. 650–657.
- [133] H. Li, Z. Lin, X. Shen, J. Brandt, and G. Hua, "A convolutional neural network cascade for face detection," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2015, pp. 5325–5334.
- [134] H. Mliki, S. Dammak, and E. Fendri, "An improved multi-scale face detection using convolutional neural network," *Signal, Image and Video Processing*, vol. 14, no. 7, pp. 1345–1353, 2020.
- [135] M. Matsugu, K. Mori, Y. Mitari, and Y. Kaneda, "Subject independent facial expression recognition with robust face detection using a convolutional neural network," *Neural Networks*, vol. 16, no. 5-6, pp. 555–559, 2003.
- [136] D. Canedo and A. J. Neves, "Facial expression recognition using computer vision: A systematic review," *Applied Sciences*, vol. 9, no. 21, p. 4678, 2019.
- [137] Y. Li, J. Zeng, S. Shan, and X. Chen, "Occlusion aware facial expression recognition using cnn with attention mechanism," *IEEE Transactions on Image Processing*, vol. 28, no. 5, pp. 2439–2450, 2018.
- [138] S. S. Farfade, M. J. Saberian, and L.-J. Li, "Multi-view face detection using deep convolutional neural networks," in *Proceedings of the 5th ACM on International Conference on Multimedia Retrieval*, 2015, pp. 643–650.

- [139] M. Ma and J. Wang, “Multi-view face detection and landmark localization based on mtcnn,” in *2018 Chinese Automation Congress (CAC)*. IEEE, 2018, pp. 4200–4205.
- [140] R. Mohammadian, M. Mahlouji, and A. Shahidinejad, “Multi-view face detection in open environments using gabor features and neural networks,” *Journal of AI and Data Mining*, vol. 8, no. 4, pp. 461–470, 2020.
- [141] D. K. Jain *et al.*, “An evaluation of deep learning based object detection strategies for threat object detection in baggage security imagery,” *pattern recognition letters*, vol. 120, pp. 112–119, 2019.
- [142] K. R. Talha, K. Bandapadya, and M. M. Khan, “Violence detection using computer vision approaches,” in *2022 IEEE World AI IoT Congress (AIIoT)*. IEEE, 2022, pp. 544–550.
- [143] E. Rahimikia and S.-H. Poon, “Machine learning for realised volatility forecasting,” *Available at SSRN*, vol. 3707796, 2020.
- [144] B. M. Pavlyshenko, “Machine-learning models for sales time series forecasting,” *Data*, vol. 4, no. 1, p. 15, 2019.
- [145] X. Wei, L. Zhang, H.-Q. Yang, L. Zhang, and Y.-P. Yao, “Machine learning for pore-water pressure time-series prediction: Application of recurrent neural networks,” *Geoscience Frontiers*, vol. 12, no. 1, pp. 453–467, 2021.
- [146] V. Chaurasia and S. Pal, “Application of machine learning time series analysis for prediction covid-19 pandemic,” *Research on Biomedical Engineering*, pp. 1–13, 2020.
- [147] Y. Yu, X. Si, C. Hu, and J. Zhang, “A review of recurrent neural networks: Lstm cells and network architectures,” *Neural Computation*, vol. 31, no. 7, pp. 1235–1270, 2019.
- [148] F. Ramezani, S. Parvez, J. P. Fix, A. Battaglin, S. Whyte, N. J. Borys, and B. M. Whitaker, “Automatic detection of multilayer hexagonal boron nitride in optical images using deep learning-based computer vision,” *Scientific Reports*, vol. 13, no. 1, p. 1595, 2023.
- [149] Y. Yu, X. Si, C. Hu, and J. Zhang, “A review of recurrent neural networks: Lstm cells and network architectures,” *Neural computation*, vol. 31, no. 7, pp. 1235–1270, 2019.
- [150] T. Young, D. Hazarika, S. Poria, and E. Cambria, “Recent trends in deep learning based natural language processing,” *IEEE Computational Intelligence Magazine*, vol. 13, no. 3, pp. 55–75, 2018.
- [151] S. Bai, J. Z. Kolter, and V. Koltun, “An empirical evaluation of generic convolutional and recurrent networks for sequence modeling,” *arXiv preprint arXiv:1803.01271*, 2018.

- [152] H. Sak, A. Senior, K. Rao, and F. Beaufays, “Fast and accurate recurrent neural network acoustic models for speech recognition,” *arXiv preprint arXiv:1507.06947*, 2015.
- [153] A. Amberkar, P. Awasarmol, G. Deshmukh, and P. Dave, “Speech recognition using recurrent neural networks,” in *2018 international conference on current trends towards converging technologies (ICCTCT)*. IEEE, 2018, pp. 1–4.
- [154] S. Ebrahimi Kahou, V. Michalski, K. Konda, R. Memisevic, and C. Pal, “Recurrent neural networks for emotion recognition in video,” in *Proceedings of the 2015 ACM on international conference on multimodal interaction*, 2015, pp. 467–474.
- [155] D. Salinas, V. Flunkert, J. Gasthaus, and T. Januschowski, “Deepar: Probabilistic forecasting with autoregressive recurrent networks,” *International Journal of Forecasting*, vol. 36, no. 3, pp. 1181–1191, 2020.
- [156] B. Lim, S. Zohren, and S. Roberts, “Recurrent neural filters: Learning independent bayesian filtering steps for time series prediction,” in *2020 International Joint Conference on Neural Networks (IJCNN)*. IEEE, 2020, pp. 1–8.
- [157] M. Schuster and K. Paliwal, “Bidirectional recurrent neural networks,” *IEEE Transactions on Signal Processing*, vol. 45, no. 11, pp. 2673–2681, 1997.
- [158] M. Liang and X. Hu, “Recurrent convolutional neural network for object recognition,” in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2015, pp. 3367–3375.
- [159] H. Salehinejad, S. Sankar, J. Barfett, E. Colak, and S. Valaee, “Recent advances in recurrent neural networks,” *arXiv preprint arXiv:1801.01078*, 2017.
- [160] S. Hochreiter and J. Schmidhuber, “Long Short-Term Memory,” *Neural Computation*, vol. 9, no. 8, pp. 1735–1780, 11 1997.
- [161] Q. V. Le, N. Jaitly, and G. E. Hinton, “A simple way to initialize recurrent networks of rectified linear units,” *arXiv preprint arXiv:1504.00941*, 2015.
- [162] K. Smagulova and A. P. James, “A survey on lstm memristive neural network architectures and applications,” *The European Physical Journal Special Topics*, vol. 228, no. 10, pp. 2313–2324, 2019.
- [163] Z. Huang, W. Xu, and K. Yu, “Bidirectional lstm-crf models for sequence tagging,” *arXiv preprint arXiv:1508.01991*, 2015.
- [164] H. Xue, D. Q. Huynh, and M. Reynolds, “Ss-lstm: A hierarchical lstm model for pedestrian trajectory prediction,” in *2018 IEEE Winter Conference on Applications of Computer Vision (WACV)*. IEEE, 2018, pp. 1186–1194.

- [165] X. Shi, Z. Chen, H. Wang, D.-Y. Yeung, W.-K. Wong, and W.-c. Woo, “Convolutional lstm network: A machine learning approach for precipitation nowcasting,” *Advances in neural information processing systems*, vol. 28, 2015.
- [166] National Institute of Standards and Technology (NIST), “NIST/SEMATECH e-handbook of statistical methods 1.3.3.1: Autocorrelation plot,” 2012, <https://www.itl.nist.gov/div898/handbook/eda/section3/autocopl.htm>, Accessed on June 26, 2023. [Online]. Available: <https://www.itl.nist.gov/div898/handbook/eda/section3/autocopl.htm>
- [167] P. Blake, E. Hill, A. Castro Neto, K. Novoselov, D. Jiang, R. Yang, T. Booth, and A. Geim, “Making graphene visible,” *Applied physics letters*, vol. 91, no. 6, p. 063124, 2007.
- [168] M. Benameur, B. Radisavljevic, J. Héron, S. Sahoo, H. Berger, and A. Kis, “Visibility of dichalcogenide nanolayers,” *Nanotechnology*, vol. 22, no. 12, p. 125706, 2011.
- [169] Z. Ni, H. Wang, J. Kasim, H. Fan, T. Yu, Y. H. Wu, Y. Feng, and Z. Shen, “Graphene thickness determination using reflection and contrast spectroscopy,” *Nano letters*, vol. 7, no. 9, pp. 2758–2763, 2007.
- [170] D. Golla, K. Chattrakun, K. Watanabe, T. Taniguchi, B. J. LeRoy, and A. Sandhu, “Optical thickness determination of hexagonal boron nitride flakes,” *Applied Physics Letters*, vol. 102, no. 16, p. 161906, 2013.
- [171] C. M. Nolen, G. Denina, D. Teweldebrhan, B. Bhanu, and A. A. Balandin, “High-throughput large-area automated identification and quality control of graphene and few-layer graphene films,” *Acs Nano*, vol. 5, no. 2, pp. 914–922, 2011.
- [172] H. Li, G. Lu, Z. Yin, Q. He, H. Li, Q. Zhang, and H. Zhang, “Optical identification of single-and few-layer mos2 sheets,” *Small*, vol. 8, no. 5, pp. 682–686, 2012.
- [173] H. Zhang, F. Ran, X. Shi, X. Fang, S. Wu, Y. Liu, X. Zheng, P. Yang, Y. Liu, L. Wang *et al.*, “Optical thickness identification of transition metal dichalcogenide nanosheets on transparent substrates,” *Nanotechnology*, vol. 28, no. 16, p. 164001, 2017.
- [174] H. Li, J. Wu, X. Huang, G. Lu, J. Yang, X. Lu, Q. Xiong, and H. Zhang, “Rapid and reliable thickness identification of two-dimensional nanosheets using optical microscopy,” *ACS nano*, vol. 7, no. 11, pp. 10 344–10 353, 2013.
- [175] B. S. Jessen, P. R. Whelan, D. M. Mackenzie, B. Luo, J. D. Thomsen, L. Gammelgaard, T. J. Booth, and P. Bøggild, “Quantitative optical mapping of two-dimensional materials,” *Scientific reports*, vol. 8, no. 1, p. 6381, 2018.

- [176] S. Balasubramaniyan, F. Parmentier, P. Roulleau, M. Thevenin, A. Brenes, and M. Trocan, “Detection of monolayer graphene,” in *Computational Collective Intelligence: 13th International Conference, ICCCI 2021, Rhodes, Greece, September 29–October 1, 2021, Proceedings 13*. Springer, 2021, pp. 784–791.
- [177] S. Balasubramaniyan, M. Thevenin, F. Amiel, and M. Trocan, “Graphene detection system,” in *Proceedings of the 14th International Conference on Management of Digital EcoSystems*, ser. MEDES '22. New York, NY, USA: Association for Computing Machinery, 2022, p. 64–69. [Online]. Available: <https://doi.org/10.1145/3508397.3564850>
- [178] K. Kang, K.-H. Lee, Y. Han, H. Gao, S. Xie, D. A. Muller, and J. Park, “Layer-by-layer assembly of two-dimensional materials into wafer-scale heterostructures,” *Nature*, vol. 550, no. 7675, pp. 229–233, 2017.
- [179] S. Masubuchi, M. Morimoto, S. Morikawa, M. Onodera, Y. Asakawa, K. Watanabe, T. Taniguchi, and T. Machida, “Autonomous robotic searching and assembly of two-dimensional crystals to build van der waals superlattices,” *Nature communications*, vol. 9, no. 1, pp. 1–12, 2018.
- [180] W. W. Dickinson, H. V. Kumar, D. H. Adamson, and H. C. Schniepp, “High-throughput optical thickness and size characterization of 2d materials,” *Nanoscale*, vol. 10, no. 30, pp. 14 441–14 447, 2018.
- [181] D. Nečas and P. Klapetek, “Gwyddion: an open-source software for spm data analysis,” *Open Physics*, vol. 10, no. 1, pp. 181–188, 2012.
- [182] J. Lee, S. Cho, S. Park, H. Bae, M. Noh, B. Kim, C. In, S. Yang, S. Lee, S. Y. Seo *et al.*, “Highly efficient computer algorithm for identifying layer thickness of atomically thin 2d materials,” *Journal of Physics D: Applied Physics*, vol. 51, no. 11, p. 11LT03, 2018.
- [183] J. Lei, J. Qu, P. Wang, H. Jiang, H. Shi, X. Sun, and B. Gao, “A facile method for precise layer number identification of two-dimensional materials through optical images,” *Optics Communications*, vol. 440, pp. 21–25, 2019.
- [184] X. Hu, C. Qiu, and D. Liu, “Rapid thin-layer ws 2 detection based on monochromatic illumination photographs,” *Nano Research*, vol. 14, pp. 840–845, 2021.
- [185] W. H. Cho, J. Shin, Y. D. Kim, and G. J. Jung, “Pixel-wise classification in graphene-detection with tree-based machine learning algorithms,” *Machine Learning: Science and Technology*, vol. 3, no. 4, p. 045029, 2022.
- [186] T. Chen, T. He, M. Benesty, V. Khotilovich, Y. Tang, H. Cho, K. Chen, R. Mitchell, I. Cano, T. Zhou *et al.*, “Xgboost: extreme gradient boosting,” *R package version 0.4-2*, vol. 1, no. 4, pp. 1–4, 2015.

- [187] T. Chen and C. Guestrin, “Xgboost: A scalable tree boosting system,” in *Proceedings of the 22nd acm sigkdd international conference on knowledge discovery and data mining*, 2016, pp. 785–794.
- [188] G. Ke, Q. Meng, T. Finley, T. Wang, W. Chen, W. Ma, Q. Ye, and T.-Y. Liu, “Lightgbm: A highly efficient gradient boosting decision tree,” *Advances in neural information processing systems*, vol. 30, 2017.
- [189] S. H. Aleithan and D. Mahmoud-Ghoneim, “Toward automated classification of monolayer versus few-layer nanomaterials using texture analysis and neural networks,” *Scientific Reports*, vol. 10, no. 1, p. 20663, 2020.
- [190] R. M. Haralick, “Statistical and structural approaches to texture,” *Proceedings of the IEEE*, vol. 67, no. 5, pp. 786–804, 1979.
- [191] R. M. Haralick, K. Shanmugam, and I. H. Dinstein, “Textural features for image classification,” *IEEE Transactions on systems, man, and cybernetics*, no. 6, pp. 610–621, 1973.
- [192] D. Mahmoud-Ghoneim, G. Toussaint, J.-M. Constans, and J. D. de Certaines, “Three dimensional texture analysis in mri: a preliminary evaluation in gliomas,” *Magnetic resonance imaging*, vol. 21, no. 9, pp. 983–987, 2003.
- [193] X. Xie, “A review of recent advances in surface defect detection using texture analysis techniques,” *ELCVIA: electronic letters on computer vision and image analysis*, pp. 1–22, 2008.
- [194] M. A. Nogueira, P. H. Abreu, P. Martins, P. Machado, H. Duarte, and J. Santos, “Image descriptors in radiology images: a systematic review,” *Artificial Intelligence Review*, vol. 47, pp. 531–559, 2017.
- [195] A. Amin and D. Mahmoud-Ghoneim, “Zizyphus spina-christi protects against carbon tetrachloride-induced liver fibrosis in rats,” *Food and Chemical Toxicology*, vol. 47, no. 8, pp. 2111–2119, 2009.
- [196] L. Moraru, S. Moldovanu, A.-L. Culea-Florescu, D. Bibicu, A. S. Ashour, and N. Dey, “Texture analysis of parasitological liver fibrosis images,” *Microscopy research and technique*, vol. 80, no. 8, pp. 862–869, 2017.
- [197] S. H. Aleithan, *Mono-to-few Layers Transition Metal Dichalcogenides, Exciton Dynamics, and Versatile Growth of Naturally Formed Contacted Devices*. Ohio University, 2018.
- [198] S. H. Aleithan, T. E. Wickramasinghe, M. Lindquist, S. Khadka, and E. Stinaff, “Growth of complex 2d material-based structures with naturally formed contacts,” *ACS omega*, vol. 4, no. 5, pp. 9557–9562, 2019.

- [199] M. M. Galloway, “Texture analysis using gray level run lengths,” *Computer graphics and image processing*, vol. 4, no. 2, pp. 172–179, 1975.
- [200] D. Swets and J. Weng, “Using discriminant eigenfeatures for image retrieval,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 18, no. 8, pp. 831–836, 1996.
- [201] E. Greplova, C. Gold, B. Kratochwil, T. Davatz, R. Pisoni, A. Kurzmann, P. Rickhaus, M. H. Fischer, T. Ihn, and S. D. Huber, “Fully automated identification of two-dimensional material samples,” *Physical Review Applied*, vol. 13, no. 6, p. 064017, 2020.
- [202] B. Wu, L. Wang, and Z. Gao, “A two-dimensional material recognition image algorithm based on deep learning,” in *2019 International Conference on Information Technology and Computer Application (ITCA)*. IEEE, 2019, pp. 247–252.
- [203] V. Badrinarayanan, A. Kendall, and R. Cipolla, “Segnet: A deep convolutional encoder-decoder architecture for image segmentation,” *IEEE transactions on pattern analysis and machine intelligence*, vol. 39, no. 12, pp. 2481–2495, 2017.
- [204] Y. Saito, K. Shin, K. Terayama, S. Desai, M. Onga, Y. Nakagawa, Y. M. Itahashi, Y. Iwasa, M. Yamada, and K. Tsuda, “Deep-learning-based quality filtering of mechanically exfoliated 2d crystals,” *npj Computational Materials*, vol. 5, no. 1, p. 124, 2019.
- [205] O. Ronneberger, P. Fischer, and T. Brox, “U-net: Convolutional networks for biomedical image segmentation,” in *Medical Image Computing and Computer-Assisted Intervention—MICCAI 2015: 18th International Conference, Munich, Germany, October 5–9, 2015, Proceedings, Part III 18*. Springer, 2015, pp. 234–241.
- [206] H.-Y. Siao, S. Qi, Z. Ding, C.-Y. Lin, Y.-C. Hsieh, and T.-M. Chen, “Machine learning-based automatic graphene detection with color correction for optical microscope images,” *arXiv preprint arXiv:2103.13495*, 2021.
- [207] K. Simonyan and A. Zisserman, “Very deep convolutional networks for large-scale image recognition,” *arXiv preprint arXiv:1409.1556*, 2014.
- [208] T.-Y. Lin, M. Maire, S. Belongie, J. Hays, P. Perona, D. Ramanan, P. Dollár, and C. L. Zitnick, “Microsoft coco: Common objects in context,” in *European conference on computer vision*. Springer, 2014, pp. 740–755.
- [209] J. Sanchez-Juarez, M. Granados-Baez, A. A. Aguilar-Lasserre, and J. Cardenas, “Automated system for the detection of 2d materials using digital image processing and deep learning,” *Optical Materials Express*, vol. 12, no. 5, pp. 1856–1868, 2022.

- [210] A. Krizhevsky, I. Sutskever, and G. E. Hinton, “Imagenet classification with deep convolutional neural networks,” *Communications of the ACM*, vol. 60, no. 6, pp. 84–90, 2017.
- [211] X. Dong, H. Li, Y. Yan, H. Cheng, H. X. Zhang, Y. Zhang, T. D. Le, K. Wang, J. Dong, M. Jakobi *et al.*, “Deep-learning-based microscopic imagery classification, segmentation, and detection for the identification of 2d semiconductors,” *Advanced Theory and Simulations*, vol. 5, no. 9, p. 2200140, 2022.
- [212] G. Huang, Z. Liu, L. Van Der Maaten, and K. Q. Weinberger, “Densely connected convolutional networks,” in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2017, pp. 4700–4708.
- [213] L. Zhu, J. Tang, B. Li, T. Hou, Y. Zhu, J. Zhou, Z. Wang, X. Zhu, Z. Yao, X. Cui *et al.*, “Artificial neuron networks enabled identification and characterizations of 2d materials and van der waals heterostructures,” *ACS nano*, vol. 16, no. 2, pp. 2721–2729, 2022.
- [214] Z. Zhou, M. M. Rahman Siddiquee, N. Tajbakhsh, and J. Liang, “Unet++: A nested u-net architecture for medical image segmentation,” in *Deep Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision Support: 4th International Workshop, DLMIA 2018, and 8th International Workshop, ML-CDS 2018, Held in Conjunction with MICCAI 2018, Granada, Spain, September 20, 2018, Proceedings 4*. Springer, 2018, pp. 3–11.
- [215] O. Oktay, J. Schlemper, L. L. Folgoc, M. Lee, M. Heinrich, K. Misawa, K. Mori, S. McDonagh, N. Y. Hammerla, B. Kainz *et al.*, “Attention u-net: Learning where to look for the pancreas,” *arXiv preprint arXiv:1804.03999*, 2018.
- [216] B. Yang, M. Wu, and W. Teizer, “Modified unet++ with attention gate for graphene identification by optical microscopy,” *Carbon*, vol. 195, pp. 246–252, 2022.
- [217] E. Ertekin, S. Tawfick, D. Adams, I. Ramos, J. Schiller, and A. Shah, “Automated image segmentation of scanning electron microscopy images of graphene using u-net neural network,” *Available at SSRN 4373023*.
- [218] S. Mahjoubi, F. Ye, Y. Bao, W. Meng, and X. Zhang, “Identification and classification of exfoliated graphene flakes from microscopy images using a hierarchical deep convolutional neural network,” *Engineering Applications of Artificial Intelligence*, vol. 119, p. 105743, 2023.
- [219] J. Wang, K. Sun, T. Cheng, B. Jiang, C. Deng, Y. Zhao, D. Liu, Y. Mu, M. Tan, X. Wang *et al.*, “Deep high-resolution representation learning for visual recognition,” *IEEE transactions on pattern analysis and machine intelligence*, vol. 43, no. 10, pp. 3349–3364, 2020.

- [220] Y. Yuan, X. Chen, and J. Wang, "Object-contextual representations for semantic segmentation," in *Computer Vision–ECCV 2020: 16th European Conference, Glasgow, UK, August 23–28, 2020, Proceedings, Part VI 16*. Springer, 2020, pp. 173–190.
- [221] X. Dong, H. Li, Z. Jiang, T. Grunleitner, I. Guler, J. Dong, K. Wang, M. H. Kohler, M. Jakobi, B. H. Menze *et al.*, "3d deep learning enables accurate layer mapping of 2d materials," *ACS nano*, vol. 15, no. 2, pp. 3139–3151, 2021.
- [222] K. P. Murphy, *Machine learning: a probabilistic perspective*. MIT press, 2012.
- [223] T. S. Ferguson, "A bayesian analysis of some nonparametric problems," *The annals of statistics*, pp. 209–230, 1973.
- [224] 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 Materials*, vol. 6, no. 3, p. 035043, 2019.
- [225] U. Von Luxburg, "A tutorial on spectral clustering," *Statistics and computing*, vol. 17, pp. 395–416, 2007.
- [226] R. M. Sterbentz, K. L. Haley, and J. O. Island, "Universal image segmentation for optical identification of 2d materials," *Scientific reports*, vol. 11, no. 1, p. 5808, 2021.
- [227] D. Comaniciu and P. Meer, "Mean shift: A robust approach toward feature space analysis," *IEEE Transactions on pattern analysis and machine intelligence*, vol. 24, no. 5, pp. 603–619, 2002.
- [228] Y.-m. Zhou, S.-y. Jiang, and M.-l. Yin, "A region-based image segmentation method with mean-shift clustering algorithm," in *2008 Fifth International Conference on Fuzzy Systems and Knowledge Discovery*, vol. 2. IEEE, 2008, pp. 366–370.
- [229] H. Zhou, X. Wang, and G. Schaefer, "Mean shift and its application in image segmentation," *Innovations in intelligent image analysis*, pp. 291–312, 2011.
- [230] M. Ester, H.-P. Kriegel, J. Sander, X. Xu *et al.*, "A density-based algorithm for discovering clusters in large spatial databases with noise." in *kdd*, vol. 96, no. 34, 1996, pp. 226–231.
- [231] R. Xu and D. Wunsch, "Survey of clustering algorithms," *IEEE Transactions on neural networks*, vol. 16, no. 3, pp. 645–678, 2005.
- [232] L. Gupta and T. Sortrakul, "A gaussian-mixture-based image segmentation algorithm," *Pattern recognition*, vol. 31, no. 3, pp. 315–325, 1998.
- [233] T. M. Nguyen and Q. J. Wu, "Dirichlet gaussian mixture model: Application to image segmentation," *Image and Vision Computing*, vol. 29, no. 12, pp. 818–828, 2011.

- [234] H. Permuter, J. Francos, and I. Jermyn, “A study of gaussian mixture models of color and texture features for image classification and segmentation,” *Pattern recognition*, vol. 39, no. 4, pp. 695–706, 2006.
- [235] H. L. Ribeiro and A. Gonzaga, “Hand image segmentation in video sequence by gmm: a comparative analysis,” in *2006 19th Brazilian Symposium on Computer Graphics and Image Processing*. IEEE, 2006, pp. 357–364.
- [236] D. H. H. Santosh, P. Venkatesh, P. Poornesh, L. N. Rao, and N. A. Kumar, “Tracking multiple moving objects using gaussian mixture model,” *International Journal of Soft Computing and Engineering (IJSCE)*, vol. 3, no. 2, pp. 114–119, 2013.
- [237] C. Shorten and T. M. Khoshgoftaar, “A survey on image data augmentation for deep learning,” *Journal of big data*, vol. 6, no. 1, pp. 1–48, 2019.
- [238] X. Li, Y. Grandvalet, F. Davoine, J. Cheng, Y. Cui, H. Zhang, S. Belongie, Y.-H. Tsai, and M.-H. Yang, “Transfer learning in computer vision tasks: Remember where you come from,” *Image and Vision Computing*, vol. 93, p. 103853, 2020.
- [239] K. Chen, J. Wang, J. Pang, Y. Cao, Y. Xiong, X. Li, S. Sun, W. Feng, Z. Liu, J. Xu, Z. Zhang, D. Cheng, C. Zhu, T. Cheng, Q. Zhao, B. Li, X. Lu, R. Zhu, Y. Wu, J. Dai, J. Wang, J. Shi, W. Ouyang, C. C. Loy, and D. Lin, “MMDetection: Open mmlab detection toolbox and benchmark,” *arXiv preprint arXiv:1906.07155*, 2019.
- [240] R. Padilla, S. L. Netto, and E. A. Da Silva, “A survey on performance metrics for object-detection algorithms,” in *2020 international conference on systems, signals and image processing (IWSSIP)*. IEEE, 2020, pp. 237–242.
- [241] E. Zhang and Y. Zhang, “Average precision.” 2009.
- [242] A. E.-J. Lim, J. Song, Q. Fang, C. Li, X. Tu, N. Duan, K. K. Chen, R. P.-C. Tern, and T.-Y. Liow, “Review of silicon photonics foundry efforts,” *IEEE Journal of Selected Topics in Quantum Electronics*, vol. 20, no. 4, pp. 405–416, 2013.
- [243] J. S. Ross, P. Klement, A. M. Jones, N. J. Ghimire, J. Yan, D. Mandrus, T. Taniguchi, K. Watanabe, K. Kitamura, W. Yao *et al.*, “Electrically tunable excitonic light-emitting diodes based on monolayer wse₂ p–n junctions,” *Nature nanotechnology*, vol. 9, no. 4, pp. 268–272, 2014.
- [244] O. Iff, D. Tedeschi, J. Martín-Sánchez, M. Moczala-Dusanowska, S. Tongay, K. Yumigeta, J. Taboada-Gutiérrez, M. Savaresi, A. Rastelli, P. Alonso-González *et al.*, “Strain-tunable single photon sources in wse₂ monolayers,” *Nano Letters*, vol. 19, no. 10, pp. 6931–6936, 2019.
- [245] C. L. Degen, F. Reinhard, and P. Cappellaro, “Quantum sensing,” *Reviews of modern physics*, vol. 89, no. 3, p. 035002, 2017.

- [246] G. Moody, V. J. Sorger, D. J. Blumenthal, P. W. Juodawlkis, W. Loh, C. Sorace-Agaskar, A. E. Jones, K. C. Balram, J. C. Matthews, A. Laing *et al.*, “2022 roadmap on integrated quantum photonics,” *Journal of Physics: Photonics*, vol. 4, no. 1, p. 012501, 2022.
- [247] P. Tonndorf, R. Schmidt, R. Schneider, J. Kern, M. Buscema, G. A. Steele, A. Castellanos-Gomez, H. S. van der Zant, S. M. de Vasconcellos, and R. Bratschitsch, “Single-photon emission from localized excitons in an atomically thin semiconductor,” *Optica*, vol. 2, no. 4, pp. 347–352, 2015.
- [248] A. Srivastava, M. Sidler, A. V. Allain, D. S. Lembke, A. Kis, and A. Imamoglu, “Optically active quantum dots in monolayer wse₂,” *Nature nanotechnology*, vol. 10, no. 6, pp. 491–496, 2015.
- [249] M. Koperski, K. Nogajewski, A. Arora, V. Cherkez, P. Mallet, J.-Y. Veuillen, J. Marcus, P. Kossacki, and M. Potemski, “Single photon emitters in exfoliated wse₂ structures,” *Nature nanotechnology*, vol. 10, no. 6, pp. 503–506, 2015.
- [250] Y.-M. He, G. Clark, J. R. Schaibley, Y. He, M.-C. Chen, Y.-J. Wei, X. Ding, Q. Zhang, W. Yao, X. Xu *et al.*, “Single quantum emitters in monolayer semiconductors,” *Nature nanotechnology*, vol. 10, no. 6, pp. 497–502, 2015.
- [251] C. Chakraborty, L. Kinnischtzke, K. M. Goodfellow, R. Beams, and A. N. Vamivakas, “Voltage-controlled quantum light from an atomically thin semiconductor,” *Nature nanotechnology*, vol. 10, no. 6, pp. 507–511, 2015.
- [252] S. I. Azzam, K. Parto, and G. Moody, “Prospects and challenges of quantum emitters in 2d materials,” *Applied Physics Letters*, vol. 118, no. 24, p. 240502, 2021.
- [253] C. E. Stevens, H.-J. Chuang, M. R. Rosenberger, K. M. McCreary, C. K. Dass, B. T. Jonker, and J. R. Hendrickson, “Enhancing the purity of deterministically placed quantum emitters in monolayer wse₂,” *ACS nano*, vol. 16, no. 12, pp. 20 956–20 963, 2022.
- [254] O. Iff, Y.-M. He, N. Lundt, S. Stoll, V. Baumann, S. Höfling, and C. Schneider, “Substrate engineering for high-quality emission of free and localized excitons from atomic monolayers in hybrid architectures,” *Optica*, vol. 4, no. 6, pp. 669–673, 2017.
- [255] R. Chaudhary, V. Raghunathan, and K. Majumdar, “Origin of selective enhancement of sharp defect emission lines in monolayer wse₂ on rough metal substrate,” *Journal of Applied Physics*, vol. 127, no. 7, p. 073105, 2020.
- [256] C. Chakraborty, K. M. Goodfellow, S. Dhara, A. Yoshimura, V. Meunier, and A. N. Vamivakas, “Quantum-confined stark effect of individual defects in a van der waals heterostructure,” *Nano Letters*, vol. 17, no. 4, pp. 2253–2258, 2017.

- [257] G. D. Shepard, O. A. Ajayi, X. Li, X. Zhu, J. Hone, and S. Strauf, “Nanobubble induced formation of quantum emitters in monolayer semiconductors,” *2D Materials*, vol. 4, no. 2, p. 021019, 2017.
- [258] G. Z. Magda, J. Pető, G. Dobrik, C. Hwang, L. P. Biró, and L. Tapasztó, “Exfoliation of large-area transition metal chalcogenide single layers,” *Scientific reports*, vol. 5, no. 1, p. 14714, 2015.
- [259] S. B. Desai, S. R. Madhvapathy, M. Amani, D. Kiriya, M. Hettick, M. Tosun, Y. Zhou, M. Dubey, J. W. Ager III, D. Chrzan *et al.*, “Gold-mediated exfoliation of ultralarge optoelectronically-perfect monolayers,” *Advanced Materials*, vol. 28, no. 21, pp. 4053–4058, 2016.
- [260] X. Qing and Y. Niu, “Hourly day-ahead solar irradiance prediction using weather forecasts by lstm,” *Energy*, vol. 148, pp. 461–468, 2018.
- [261] S. Bouktif, A. Fiaz, A. Ouni, and M. A. Serhani, “Optimal deep learning lstm model for electric load forecasting using feature selection and genetic algorithm: Comparison with machine learning approaches,” *Energies*, vol. 11, no. 7, p. 1636, 2018.
- [262] A. G. Salman, Y. Heryadi, E. Abdurahman, and W. Suparta, “Single layer & multi-layer long short-term memory (lstm) model with intermediate variables for weather forecasting,” *Procedia Computer Science*, vol. 135, pp. 89–98, 2018.