

Investigating Convective Flow in Aluminium Solidification Using X-ray Imaging and Deep Learning



Rohit Abraham

Corpus Christi College

University of Oxford

Supervised by:

Dr. Enzo Liotti

Prof. Andrew Zisserman

A thesis submitted for the Part II examination of

Final Honour School of Materials Science

Acknowledgments

During this past year, the support of my supervisors has been paramount to the completion and enjoyment of this project. I would like to thank Dr. Enzo Liotti for investing so much time over the year towards the project and my development; it was a true pleasure to work together. I would also like to extend thanks to Prof. Andrew Zisserman, from whom it was an honour to receive guidance and advice from. I must also thank Dr. Carlos Arteta who voluntarily took time out of his schedule to guide me through various errors and bugs in my code which, at times, I thought were unfixable.

I also wish to express my appreciation towards my college tutors, Prof. David Armstrong, Prof. Peter Nellist, and Prof. Christopher Patrick, who spent the previous four years building my confidence in materials science and beyond, making this degree so much more manageable.

I would like to thank my family; without their support none of this would have been remotely possible. Finally, I would like to acknowledge the invaluable support of Ewan, Yix, Jake, Alex, Myra, and many others who's care and friendship not only made the last few years so enjoyable, but defined my time at Oxford.

Abstract

Understanding the effect of convective flow during solidification of a metal can assist in improving processing methods, leading to greater control over material properties. Previous methodologies have been yet to utilise the precision and efficiency offered by modern deep learning models to provide quantitative analysis of convective flow and resulting solidification phenomena. The thesis aims to develop and utilise a deep learning object detection model to count grain nucleation events in *in situ* synchrotron X-ray radiography sequences of Al-Cu alloys. A Faster R-CNN model was trained on a large dataset of annotated sequences and applied to 19 experiments with and without electromagnetic field stirring to induce convective flow. The new object detection model showed improved performance over a wider range of experiments and reduced analysis time compared to prior work. Analysis performed verified literature on nucleation occurring in bursts showed that forced convection led to faster average nucleation rates, combined with a reduced observation of nucleation burst behaviour. Through this process, several steps have been identified to improve the model further, and detailed what future solidification experiments need to be conducted to present a more vigorous investigation into convective flow.

Contents

1	Introduction	8
1.1	Engineering Context	8
1.2	Imaging Solidification via. In Situ X-ray Radiography	9
1.3	Project Motivations	10
1.4	Aims	11
2	Literature Review	12
2.1	Imaging Metal Solidification	12
2.2	What is Machine and Deep Learning	15
2.3	Computer Vision and Object Detection	16
2.4	Solidification in Al and Al-Cu Systems	18
2.4.1	Introduction	18
2.4.2	Wave Nature of Nucleation	19
2.4.3	Convective Effect on Solidification	21

CONTENTS

2.5 Integrating Machine Learning with Solidification Analysis	22
3 Developing an Object Detection Model	24
3.1 Faster R-CNN Architecture	25
3.1.1 Model Selection	25
3.1.2 Details of the Faster R-CNN Architecture	26
3.1.3 Technical Implementation	27
3.2 Dataset	28
3.2.1 Data Acquisition	28
3.2.2 Data Split	30
3.2.3 Data Processing Pipeline	31
3.3 Model Training and Evaluation	33
3.3.1 Training	33
3.3.2 Evaluating Object Detection Models	34
3.3.3 Measuring Performance Changes During Solidification Sequences . .	37
3.4 Model Results	38
3.4.1 Reducing the Dataset for 25 wt % Cu Alloys	39
3.4.2 Key Metrics	40
3.4.3 Precision-Recall Curves	40

CONTENTS

3.4.4	Understanding Model Predictions	41
4	Nucleation Event Analysis	44
4.1	Experimental Methods	44
4.2	Analytical Methods	45
4.2.1	Grain Nucleation Counting	45
4.2.2	Smoothing and Refining Grain Count Curves	46
4.3	Bursts of Nucleation	47
4.4	EMF Stirring	48
4.4.1	Comparing Burst Behaviour	48
4.4.2	Low Cooling Rates	50
4.4.3	High Cooling Rates	51
4.4.4	Average Nucleation Rates	51
5	Future Work	53
5.1	Improving The Grain Detection Model	53
5.1.1	Training Dataset	53
5.1.2	Overlapping Grains and Late Sequence Detection	54
5.1.3	Grain Segmentation	55
5.1.4	Future Model Architectures	56

CONTENTS

5.2	Investigating Convection	57
6	Conclusions	58
7	Project Reflections	60
7.1	Ethics and Sustainability	60
7.2	Project Management	62
7.2.1	Defining a Project Scope and Structure	62
7.2.2	Personal Development	63
7.2.3	Project Timeline	64
7.2.4	Project Management Forms	66
7.3	Health and Safety	73
A	Appendix	74
A.1	Evolution of Region-Based CNN Architectures	74
A.1.1	R-CNN	74
A.1.2	Fast R-CNN	75
References		76

Chapter 1

Introduction

1.1 Engineering Context

The global metal industry is a significant contributor to carbon emissions, with steel and aluminium production representing around 9% and 2% of global carbon emissions, respectively [1, 2]. Given the impact on the environment due to these industries, decarbonisation is a priority to reach global emission targets. Both aluminium and steel can be theoretically recycled and processed with zero detriment to final material properties, making them model materials for the circular economy [3]. Therefore, decarbonising and improving processing efficiencies within the recycling of metals is of strategic research and commercial interest [4].

Steps towards the reduction of emissions within steel and aluminium recycling has focused on redesigning equipment used, such as using carbon-free inert anodes in the electrolysis of aluminium [5], and shifting energy sources used during recycling to green alternatives.

The research presented in this thesis surrounds the metallurgical improvements which could be applied to recycling. In order to improve efficiency, the control of microstructure

can enable faster and less carbon intensive routes towards desirable products from the starting scrap metal. A significant part of microstructural control derives from the manipulation and understanding of the solidification process, which has been explored in this thesis through the study of nucleation and grain growth dynamics in an aluminium-copper system.

1.2 Imaging Solidification via. In Situ X-ray Radiography

In metallurgy, solidification is the process through which liquid metal undergoes a phase transformation to a solid state. During this process, conditions such as cooling rates, thermal gradients, and alloy composition play a critical role in determining the microstructure, mechanical properties, and the formation of defects [6].

During the early stages of solidification, crystals nucleate where the temperature exists below the equilibrium freezing point. These grains experience various forces such as fluid flow, convection or gravity, which affect the dynamics of the overall solidification. By looking at the effect of solidification conditions on grain motion and behaviour in these stages, a link can be made between those changes in conditions and wider phenomena such as microsegregation, macrosegregation and defect formation.

In order to experimentally investigate solidification processes, most research can be categorised into two approaches: a *post-mortem* analysis of the solid metal, or *in situ* imaging during the solidification process. Techniques such as optical microscopy, scanning electron microscopy and transmission electron microscopy provide imaging at various length scales and level of detail, and has been used to investigate the final effect of various changes to conditions during solidification [7–10].

While *post-mortem* analysis is crucial to understanding the final result of solidification,

it does not capture the physical processes and dynamics of microstructural changes [11].

On the other hand, *in situ* imaging enables real-time visualisation and monitoring of solidification [12], allowing one to identify individual nucleation events, or the successive growth and movement of dendrites. Solidifying metals from a liquid state requires both very high temperatures, and careful control of the surrounding environment, which limits the range of imaging techniques that can be used [11]. In recent years, *in situ* X-ray radiography has emerged as an effective method to image the nucleation and evolution of dendrites, and is the method in which the data analysed in this thesis has been collected.

Collecting large volumes of data through conducting solidification experiments introduces a common bottleneck experienced in many fields, scientific and beyond: the time required for analysing and interpreting the data limits the amount of information extracted and therefore the scientific output.

1.3 Project Motivations

Prior work by Liotti *et al.* (2018) sets the primary context for the investigation in this thesis. A machine learning (ML) model was developed to perform object detection on grains during the solidification sequences as they nucleate, allowing each frame of a sequence to have the number of grains automatically counted. By doing so, quantitative data was presented into the effect of cooling rates and thermal gradients on nucleation behaviour [13]. This project will build upon this by developing a modern deep learning (DL) approach to grain detection and use the model to present preliminary analysis on solidification experiments with forced convection through EMF stirring.

The distinction between traditional ML and DL models is a central theme of this thesis as it is the core change in methodology between the prior work carried out and the techniques used in this project. By utilising improvements in computer vision and ad-

1.4 Aims

vanced deep learning networks, a model will be developed which can perform detections on a more generalised dataset, leading to higher quality analysis on varying conditions. Improvements in efficiency will be demonstrated both in the development of the model and during its use for performing grain nucleation event analysis.

1.4 Aims

The main goals of this project are as follows:

1. To develop a generalised deep learning object detection model for aluminium grains in X-ray radiography solidification sequences.
2. To benchmark the new object detection model on grain counts with the previous machine learning model developed.
3. To perform preliminary analysis on previously unanalysed solidification experiments with convective flow induced through EMF stirring.

Chapter 2

Literature Review

In order to provide context and refine the objectives of this thesis, a brief review has been conducted on: developments in *in situ* X-ray radiography studies of aluminium alloy solidification; the recent history and improvements in object detection models; and how we can combine both of these fields to explore the effect of convective flow on the nucleation and growth of equiaxed dendrites.

2.1 Imaging Metal Solidification

X-ray sources have been employed in several methods to image materials at varying length scales, enabling different features and data to be collected. Diffraction techniques provide detail of the crystal structure, defects and phases to be measured [14]. X-ray spectroscopy uses absorption to characterise surfaces and interfaces with length scales of 1 - 500 nm, producing insights on chemical composition [15]. Transmission based X-ray imaging techniques including radiography and tomography enable one to investigate at length scales $>1\text{ }\mu\text{m}$ [11]. The nucleation and growth of crystals and defects, alongside phenomena such as microsegregation and macrosegregation can be identified.

2.1 Imaging Metal Solidification

In situ X-ray imaging, specifically radiography and tomography have been developed to become one of the most prominent methods to study metal solidification in recent years [12, 16–18]. While tomography, a three-dimensional approach to imaging, naturally gathers more information about spatial arrangements of crystals and microstructure development, radiography using thin samples produces insights with reduced experimental and data complexity [11].

The development and increase in availability of synchrotron radiation sources have catalysed the aforementioned the growth and adoption of *in situ* imaging techniques. Synchrotron facilities allow for high-energy and high-brilliance x-ray sources [18], enabling high resolution imaging with clear detail in real time. This is due to the higher penetration depth of synchrotron radiation enabling thicker samples to be imaged, and the higher brilliance reducing exposure time, allowing for dynamic processes to be captured at higher frame rates [16].

Approaches towards *in situ* experimental set-ups have converged towards relatively standardised methods for either tomography both radiography. For radiography, a thin sample is prepared and enclosed within Bridgeman-style furnaces [18] with resistive or laser based heating units to melt samples and subsequently control cooling rates. The imaging rig can be customised as per research requirements including mechanical stages to perform deformation on the sample [19], and electromagnetic systems to mix samples [13]. Incident X-rays pass through the sample and reach a scintillator, which produces visible light relative to the electron density at the point passed through the sample. Lenses focus the light towards a detector which is typically CCD or CMOS based [20]. A example diagram is given in **figure 2.1**, where it should be noted the vertical nature of the furnace will introduce gravitational effects to the sample during solidification. Tomography set-ups are similar but include a cylindrical sample and a rotating stage to capture the sample in

2.1 Imaging Metal Solidification

three dimensions [18].

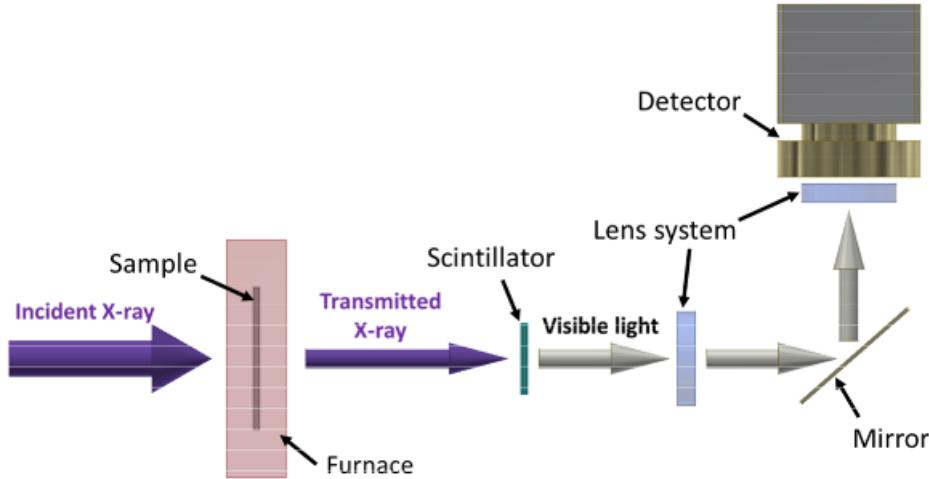


Figure 2.1: Schematic of an example experimental set up for *in situ* solidification experiments using synchrotron radiation by Feng *et al.* (2020) [20]

This experimental set up provides a system through which conditions of solidification can be carefully controlled. This includes: cooling rates, thermal gradients, solidification velocities, alloy compositions, and atmospheric control [18]. The improvements in detectors and their capture rates have also enabled imaging of faster cooling rates while maintaining an *in situ* methodology.

Collected data is usually subject to several processing steps before it can be analysed. These analytical techniques can often represent the most time intensive stage in the research pipeline. Due to the high demand of synchrotron beam time, most studies focus on conducting as many experiments as possible during their allocated time. Therefore, analysis and interpretation of results can occur between months and years after experiments were initially conducted. Machine learning techniques have been increasingly applied towards these experiments to enable rigorous and automated quantitative analysis [13, 21], reducing the time-induced bottleneck encountered during the research timeline.

2.2 What is Machine and Deep Learning

Before exploring how machine learning can be applied to solidification analysis, a brief section is provided on the background of ML and DL developments. Machine learning is where algorithms are used to automatically discover patterns in data, providing insights such as classification or predictions [22]. A machine learning model learns from existing data (training set) to adjust their parameters (weights), so that the patterns specific to a dataset are not explicitly programmed into the algorithm, but rather are learnt.

ML tasks can be often categorised as either classification problems, where the model outputs a discrete label (e.g. categorising images of animals), or a regression problem, where the model outputs a continuous value (e.g. predicting energy demand) [22]. Traditional ML models rely on feature engineering, where the method in which information and patterns (features) are extracted is inherently hard coded into the algorithm, which represents a time intensive part of model development.

Deep learning is a subset of machine learning which uses artificial neural networks which consist of many interconnected neurons that automatically learn features during the training process from much larger amounts of data. DL models, such as deep convolutional neural networks (CNNs), have a significantly larger number of parameters resulting in more robust models which can perform well with varying complex datasets [23].

The growth of DL has been enabled by both algorithmic developments alongside improvements in hardware, specifically graphical processing units (GPUs). GPUs can perform parallel calculations with significantly more cores compared to a central processing unit (CPU), reducing the training and inference time for models with millions, and more recently 100s of billions, of parameters [24]. **Figure 2.2** demonstrates the increase in computational resources required to train various ML and DL models, highlighting the surge

2.3 Computer Vision and Object Detection

in 2012 when deep CNNs became more prominent.

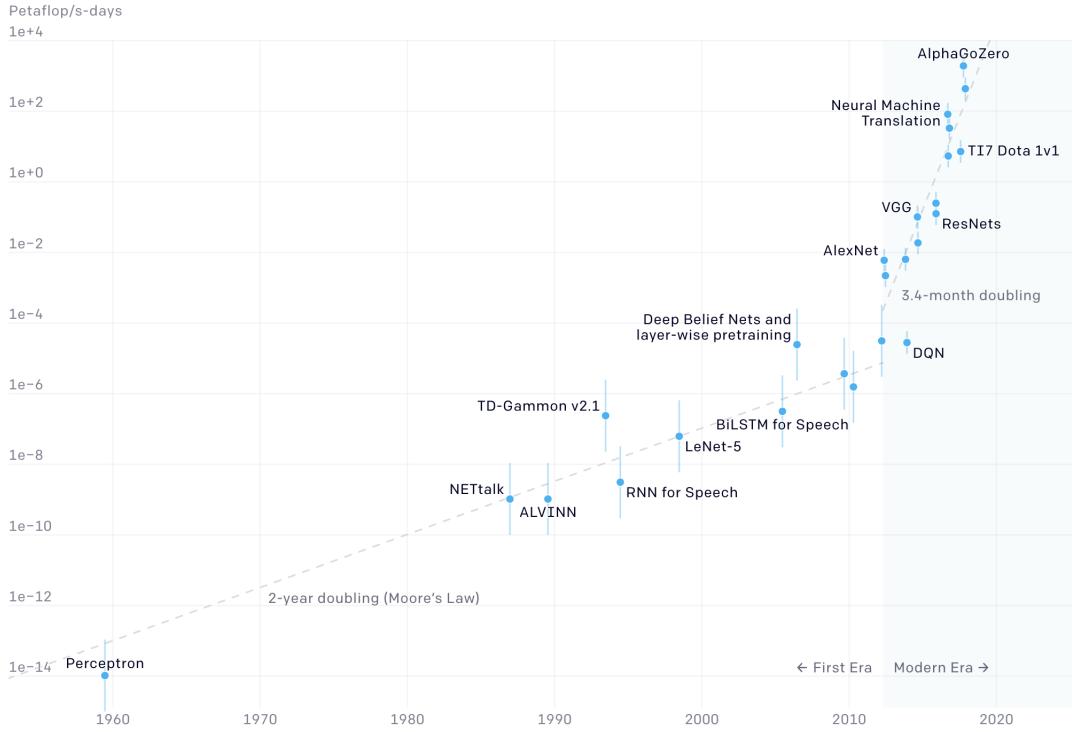


Figure 2.2: Log-scale of total compute (Petaflops/day) used to train a selection of prominent models over the last 60 years. Figure is by Amodei *et al.* on behalf of OpenAI, last updated in 2023 [25].

2.3 Computer Vision and Object Detection

Computer vision is a field within ML surrounding the extraction of useful data from images or videos. Classification is a task within the field of computer vision where a model aims to label entire images based on its contents, for example labelling pictures of various fruits from a set list of classes. Object detection is where objects within an image are located via bounding boxes, and then each region is classified with a label. Since the coordinates of the bounding box are a set of continuous variables, object detection represents both a classification and regression task.

Approached towards object detection initially consisted of customised feature extraction, such as the Viola-Jones detector first proposed in 2001, which used series of cascading

2.3 Computer Vision and Object Detection

classifiers to efficiently and accurately detect faces [26]. This was followed by sliding window approaches, where a classifier scans over an image to find and label objects, and became a popular technique. In 2005, Dalal and Triggs proposed the Histogram of Oriented Gradients (HOG) approach, where histograms were used to produce localised representations of gradients that show the structure of the contents of the image [27]. The HOG approach was combined with support vector machines, a type of algorithm which identifies patterns or classifies data by finding a boundary (hyperplane) between the data. When the HOG is data represented in a multidimensional space, it will exist on a certain side of the hyperplane, identifying which category it belongs to [22].

DL approaches to object detection first began through the classification stage in the task. In 2012, Krizhevsky *et al.* first demonstrated very large, deep convolutional neural networks being applied image classification, surpassing the performance of all previous models. The CNN first extracts features through 5 convolutional layers and 3 pooling layers, and flatten them into 1D feature vectors. The feature vector then passes through fully connected layers where activation functions process the vectors and finally outputs a vector with shape of $(1, n)$ where n is the number of classes. The value of each row represents the probability the image belongs to that class [28]. In 2014, Simonyan & Zisserman introduced the VGG-16 network which made use of 16 total layers and smaller convolutional filters, which enabled more detailed and complex features to be learnt [29]. More modern CNN classifiers exhibit deeper networks with new approaches to certain techniques in model design and training. Classifiers such as ResNet have variations with 50, 101 and 152 total layers [30].

Regarding bounding box regression, a similar progression occurred with R-CNN which first used a pre-DL selective search approach to first propose regions where objects may be and then used a CNN classifier to assign each region to a class [31]. The method through

which regions are proposed moved towards neural networks which work in conjunction with the classifier, leading to models such as Faster R-CNN. Further details outlining the progression of R-CNN based object detection methods are explored in **section 3.1** and **appendix A.1**.

2.4 Solidification in Al and Al-Cu Systems

2.4.1 Introduction

During solidification, the intrinsic and extrinsic conditions of the system has substantial impact on the microstructure, mechanical properties and defect formation in a metal. By understanding and controlling these conditions, metals and alloys can be tailored towards specific applications, demonstrate better mechanical performance, and undergo failure less often [6]. When considering aluminium and aluminium alloys, the primary techniques to influence solidification include controlling the cooling rate and thermal gradients, using stirring methods, using grain refiners, employing different mold materials and coatings, among others [6, 32–34].

The aluminium-copper alloy system has a wide range of applications in industry and research. Al-Cu alloys have become a common choice for aerospace and automotive applications [35] due to high strength to weight ratio, with relatively high hardness and toughness. Commercially, Al-Cu alloys are denoted as the 2xxx series, with typical Cu content varying between 2 - 6 wt %. For example, Al 2024-T3 is alloyed with 4 % Cu and smaller percentages of Mg, Mn, Si, Ti and Zn, and is commonly used for aircraft gears, fittings, bolts, due to a high vickers hardness of 137 MPa and tensile strength of 345 MPa [36].

Higher concentrations of Cu between 10 - 25 wt % have fewer commercial uses but represent a system which can enable research in a variety of fields for a few key reasons. Firstly,

2.4 Solidification in Al and Al-Cu Systems

Al-Cu alloys are relatively easy to produce in a lab based setting, where the composition can be precisely controlled [37]. When considering *in situ* solidification imaging, the Al-Cu system exhibits strong contrast between the α -Al crystals and the liquid Al-Cu phase [38]. This allows for clearer details when analysing solidification phenomena. Additionally, at 10 wt %, it is noted that the buoyancy of α -Al crystals in the liquid phase is neutral, meaning gravitational effects are less apparent and crystal flow and movement can be naturally observed. When considering the phase diagram in **figure 2.3**, a well defined eutectic reaction exists at 33 wt % Cu. Compositions between 10 - 25 wt % are in the hypoeutectic region, which represents a reaction commonly observed in many other alloy systems of research interest [6]. Due to these reasons, Al-Cu is considered to be a model alloy system.

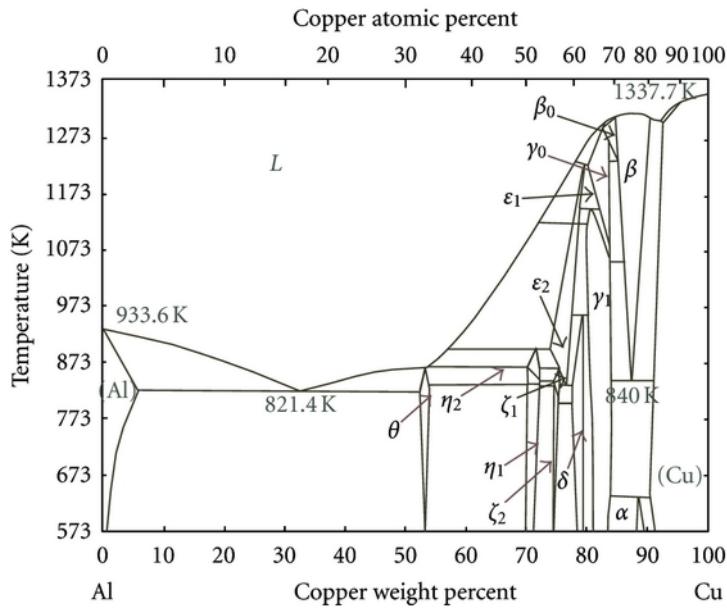


Figure 2.3: Phase diagram of the Al-Cu system by Huerta *et al.* (2012) [39]

2.4.2 Wave Nature of Nucleation

In an Al-Cu system, the two growth morphologies of equiaxed and columnar dendrites are of particular importance. In industry, equiaxed grains are commonly achieved by the use of grain refiners, such as TiB₂ and TiC [40]. This morphology is favoured due to

isotropic mechanical properties, enhanced ductility and resistance to crack propagation, while columnar dendrites are encouraged where mechanical anisotropy is desired. During solidification, high thermal gradients encourage directional grain growth while lower cooling rates enable diffusive processes to occur, resulting in long needle-like columnar grains. With low thermal gradients and faster cooling rates, smaller randomly oriented equiaxed grains are formed [6, 41].

X-ray radiography has been increasingly employed to conduct studies on both equiaxed and columnar dendrites, furthering understanding on the variation in crystal shapes, crystal growth and defect formation [18]. Imaging and analysing the early stage of nucleation of these grains represents a more difficult challenge due small size and number of crystals [13, 42], hence fewer works exist.

In 2015, Prasad *et al.* qualitatively demonstrated that at Al-Si and Al-15 wt % Cu exhibit bursts of nucleation when under a small thermal gradient [43, 44]. This effect was first outlined by StJohn *et al.* when the interdependency theory was proposed in 2011. When nucleation occurs, the local concentration of the solute increases, resulting in the equilibrium freezing temperature increasing around the recently formed crystals. Nucleation becomes locally less favourable to occur because of a decrease in the available undercooling, leading to the formation of a solute suppressed nucleation zone (SSNZ) [45]. Liotti *et al.* (2018) followed with a quantitative analysis of this effect in both Al-10 wt % Cu and Al-25 wt % C, demonstrating that bursts were more noticeable in higher solute concentrations and that presence of an SSNZ can lead to a higher total number of nucleation events through the enablement of less efficient but more populous nucleants. The paper also looked at the effects of varying composition and cooling rates on nucleation, specifically the rate of nucleation [13]. The bursts of nucleation have also been observed in other alloy systems, such as Al-Pt-Er, which was used by Feng *et al.* (2021) to model the nucle-

ation of $\text{Al}_{13}\text{Fe}_4$, an inter-metallic compound (IMC) used in industrial applications. This study demonstrated the burst phenomenon in ordered IMCs for the first time, contrasting previous studies surrounding the nucleation of disordered phases from solid solutions[42].

2.4.3 Convective Effect on Solidification

Stirring during solidification has been commonly used as a technique to induce convection in a melt, and by doing so, promote uniform temperatures and homogeneous solute distributions. This is performed to reduce the likelihood of defects forming through avoiding porosity, reducing segregation and encouraging homogeneous grain structures.

Stirring is typically achieved through two methods, mechanical and electromagnetic (EMF). Mechanical stirring involves agitating the melt either through a physical device, such as a blade, or inert gases bubbled through [46]. EMF stirring uses electromagnetic fields to induce fluid flow during solidification, preventing any physical contact with the melt and allowing for precise control over magnitude and direction of stirring. EMF stirring has shown to affect transport phenomena, thermal distribution, and final microstructure [47, 48], and is a preferred method to reduce porosity and segregation in the continuous casting of steels in industry [49–51].

Prior studies on the physical processes and resulting effects of EMF stirring on solidification have mostly consisted of numerical simulations [52, 53] and *post-mortem* imaging [54, 55]. While these studies have been able to advance understanding on the effects of stirring on final microstructure, *in situ* experiments to investigate physical dynamics of EMF stirring are less common. In 2014, Liotti *et al.* employed *in situ* X-ray radiography to develop a method of studying dendrite fragmentation in Al-15 wt % Cu through pulsed EMF. This method was used to demonstrate that dendrite fragmentation rate can be increased significantly through forced convection of the interdendritic liquid phases [56].

This was followed by a similar study by Liotti *et al* (2016) which further collected more data on the dendrite fragmentation under different conditions in this system, revealing particular stirring directions having greater effects on fragmentation rates. With the wider, reproducible study, a theoretical model of fragmentation behaviour was developed [57].

2.5 Integrating Machine Learning with Solidification Analysis

In the study by Liotti *et al.* (2018) on the wave nature of nucleation, a density estimation approach was taken via hand crafted feature extraction to perform object detection on the grains [13]. The ML approach was employed to minimise the amount manual annotation of training data required, and used the optimal technological approach available during the model's development. This approach led to vigorous quantitative results, but exhibited the previously discussed limitations of a pre-deep learning method, such as limited flexibility of the model to be applied to the imaging of different conditions. Images also required extensive pre-processing, often requiring manual testing for each experiment. Feature engineering represented a large proportion of the development timeline, and meant the final model performed well but for a limited and specific dataset.

Karagadde *et al.* used pixel-level segmentation in their 2019 study on porosity formation through *in situ* X-ray tomography (XCT) experiments of semi-solid Al alloys [58]. While the exact method used was not detailed, the ML approach was implemented using the Weka plugin for Fiji (ImageJ), and would have occurred through a variety of trainable features and classifiers available in the software [59, 60]. This is an example of where traditional ML approaches have been established for many years, and are packaged with graphical user interfaces for scientific image analysis.

More recently, work has been carried out using deep learning approaches, for example, Shashank *et al.* used CNNs to segment Al and Al₂Cu in XCT data of solidified samples. One of the notable conclusions suggested that the significant decrease in the time required for segmentation enabled by a DL approach could enable 4D analysis (tomography data with a time series) in the near future [61].

Other examples have highlighted the ability to quickly train models and deploy them in both characterisation and segmentation based tasks for variety of XCT and radiography studies, while achieving high accuracy when evaluating the model on test data [62, 63].

The recent applications of ML and DL techniques have demonstrated large improvements in automating the analysis of solidification studies in significantly shorter timeframes. The ML approaches, such as those employed by Liotti *et al.* (2018) and Karagadde *et al.* (2019) have set the groundwork for analysis in this field, while DL approaches are beginning to be realised as a method in which further accuracy, flexibility and processing speed can be achieved.

Despite this, one can acknowledge the gap in integration of DL techniques towards *in situ* solidification experiments, presenting the potential for improvements in further analytical speed and model robustness. Specifically, a DL model of this nature could be used to present quantitative results on the effects of convective flow on equiaxed dendrites induced by EMF stirring.

To summarise, while considerable progress has been made to both the experimental methods of imaging and analytical techniques for results, there exists opportunity for further research in this field. Leveraging modern DL models and the advancements in *in situ* imaging could enable deeper quantitative insights into solidification processes in aluminium alloys and other alloy systems.

Chapter 3

Developing an Object Detection Model

This section outlines the steps taken to develop the object detection model used to identify aluminium grains in X-ray radiography images. It also explores the background of any technology or scientific theory used relevant to the object detection model.

The development of a DL model typically follows a standard structure, whereby each step can be modified and adjusted to result in a suitable end model specific to the task. This section will initially cover the background of the model architecture used for this task and how it was implemented. It will then follow the progression of **figure 3.1**, covering the data used, how it was prepared, and split, followed by an exploration into how the model is trained on said data. Model evaluation methods will be explored for object detection tasks, and look at how this was used verify the performance of grain detection. The section will conclude with the final model results, evaluating its performance and a brief discussion on optimising the model for more specific analytical use-cases. In order for this section to follow the development pathway, certain techniques, theories or terminology may be

3.1 Faster R-CNN Architecture

referred to before they are thoroughly explained; where this occurs a note has been made for clarity.

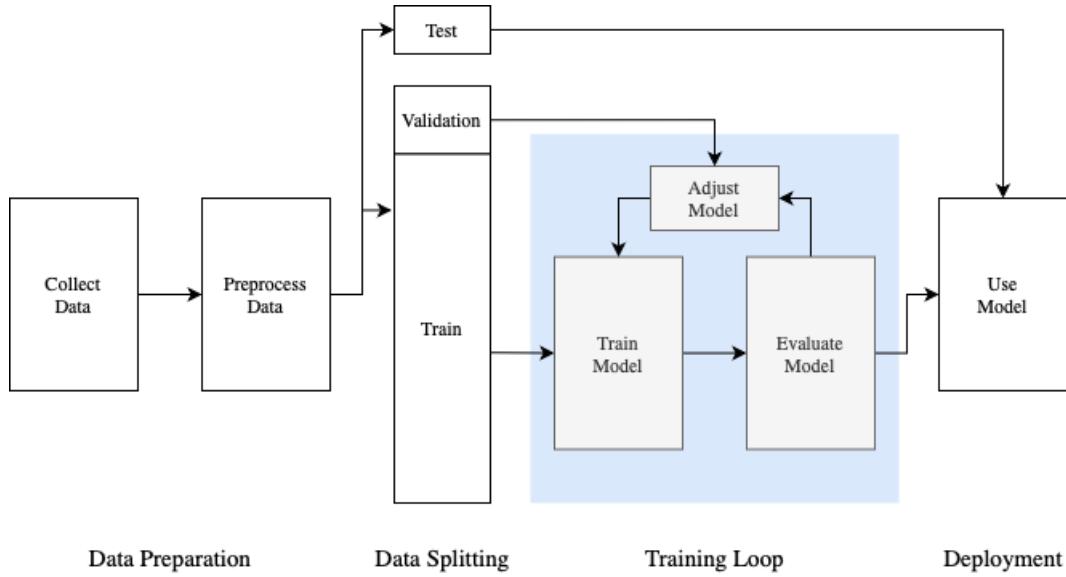


Figure 3.1: Model Development Pathway (adapted from Dragonfly Documentation (2020) [64])

3.1 Faster R-CNN Architecture

3.1.1 Model Selection

The task of detecting and classifying the grains during solidification is a standard object detection task with two classes to determine between: the grain and the background. While many advanced deep learning object detection models exist, and continue to be developed with improving precision [65], one must carefully choose the most appropriate model given project constraints and objectives. The Faster R-CNN (Region-based Convolutional Neural Network) model [66] satisfied a number of key these requirements.

Despite being initially proposed in 2015, Faster R-CNN is still a popular choice for object detection tasks. High performance and precision has been demonstrated with varying custom datasets [67, 68], object sizes and image resolutions [69]. This is partly due to the two-stage architecture which focuses on accuracy of detection, compared to single stage detectors such as YOLO (You Only Look Once) and SSD (Single Shot Detector) [70, 71].

3.1 Faster R-CNN Architecture

which prioritise speed of detections. Further advantages include a large community of support and resources available, with efficient integration with popular machine learning frameworks such as PyTorch [72], enabling relatively fast deployment with a reduced learning curve relative to newer networks.

3.1.2 Details of the Faster R-CNN Architecture

The Faster R-CNN architecture uses a two-stage approach to object detection, with a schematic presented in **figure 3.2**. A brief exploration into the evolution of region-based CNNs is given in **appendix A.1**. A convolutional neural network, such as VGG-16 or ResNet-50, produces a convolutional feature map for an input image. This CNN is referred to as the backbone, and performs the feature extraction and classification. In order to identify where objects exist within the image, Faster R-CNN introduces a region proposal network (RPN). This differs from previous techniques in which regions were proposed via pre-deep learning algorithms such as selective search [73]. As detailed by Ren *et al.* in 2015, the RPN is a fully convolutional network that takes an input image, and outputs a set of object proposals and an associated objectness score (a metric describing likelihood that proposal is an object rather than background) [66].

Using a deep learning approach towards region proposal brings about two main advantages. Firstly, the accuracy of detection is improved as the RPN is now trained on the dataset, and will dynamically make adjustments relative to the training data. This is contrasted to prior methods which proposed regions through fixed, dataset agnostic techniques. Secondly, the RPN shares the convolutional feature map which is produced by the object detection network [66, 74], which addresses the efficiency bottleneck of using selective search techniques.

Once regions are proposed, they are combined with the feature map to create regions of

3.1 Faster R-CNN Architecture

interest, formally called a Region of Interest (RoI) pooling layer. Each RoI has a respective feature vector, where the model then classifies the probability it belongs to the grain class, and outputs the associated bounding box coordinates [66, 74].

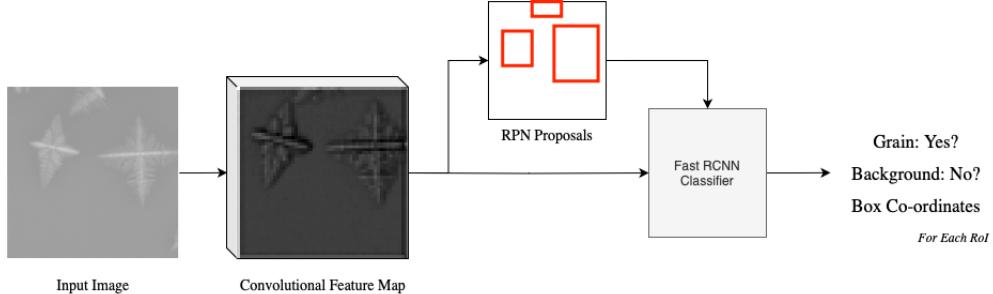


Figure 3.2: Faster R-CNN network overview (adapted from Ren *et al.* (2015) [66])

3.1.3 Technical Implementation

While machine learning, and specifically deep learning models can be developed in a range of programming languages, Python has emerged as the most popular and has been used for almost all parts of the project's model development. Most major developments in deep learning and neural networks have been implemented in Python [75], and this has been catalysed by the availability of many open-source machine learning libraries. These libraries offer a collection of scripts and functions to assist with various elements of the development pipeline, including pre-written algorithms, models with pre-trained weights, data preprocessing tools, among other features.

PyTorch [76] is a well established library which was first introduced in 2016, and been utilised by both industry and academia to develop machine and deep learning solutions for a wide range of tasks including language, audio and computer vision [77]. PyTorch was chosen as the main platform on which the object detection model was developed due to its relative ease of implementation, availability of many educational resources and support, and efficient hardware integration with GPUs made available during this project. As part of PyTorch, the torchvision [78] library was used specifically to implement computer vision

3.2 Dataset

related models and necessary tools to assist with development.

Faster R-CNN is included as one of the object detection models in torchvision, which was used with a ResNet-50-FPN backbone, the same CNN used by Ren *et al.* when benchmarking the model in 2015 [30, 66]. The torchvision repository on GitHub contained several object detection reference scripts which were used to significantly speed up development time [78]. These scripts included pre-written functions for training loops and model evaluation¹, some of which were modified to improve handling of potential errors within the dataset and associated annotations.

3.2 Dataset

Deep learning models excel by learning from large volumes of labelled training data. In this section, the acquisition of this labelled data is detailed, alongside key processing steps taken to improve and augment the quality and amount of data available. Finally, a brief outline of how the data is implemented into PyTorch and the model is included.

3.2.1 Data Acquisition

As detailed by Liotti *et al.* in 2018, the data used in this project originates from *in situ* solidification experiments of four samples of Al-Cu alloys with compositions of 10, 15, 20 and 25 wt % Cu with 0.1 wt % Ti acting as a grain refiner to encourage equiaxed grains to form. Samples were cooled from a liquid state in a purpose-built solidification rig at cooling rates of 0.3, 0.7 and 1.5 K s^{-1} . The solidification sequences were imaged via X-ray radiography from a Synchrotron source with a GGG:Eu scintillator, where light was magnified and projected to a pco.DIMAX camera, and captured at rate of 6.67Hz [13].

Example images from the dataset are shown in **figure 3.3**.

¹Further details of model evaluation techniques will be explored in **section 3.3.2**

3.2 Dataset

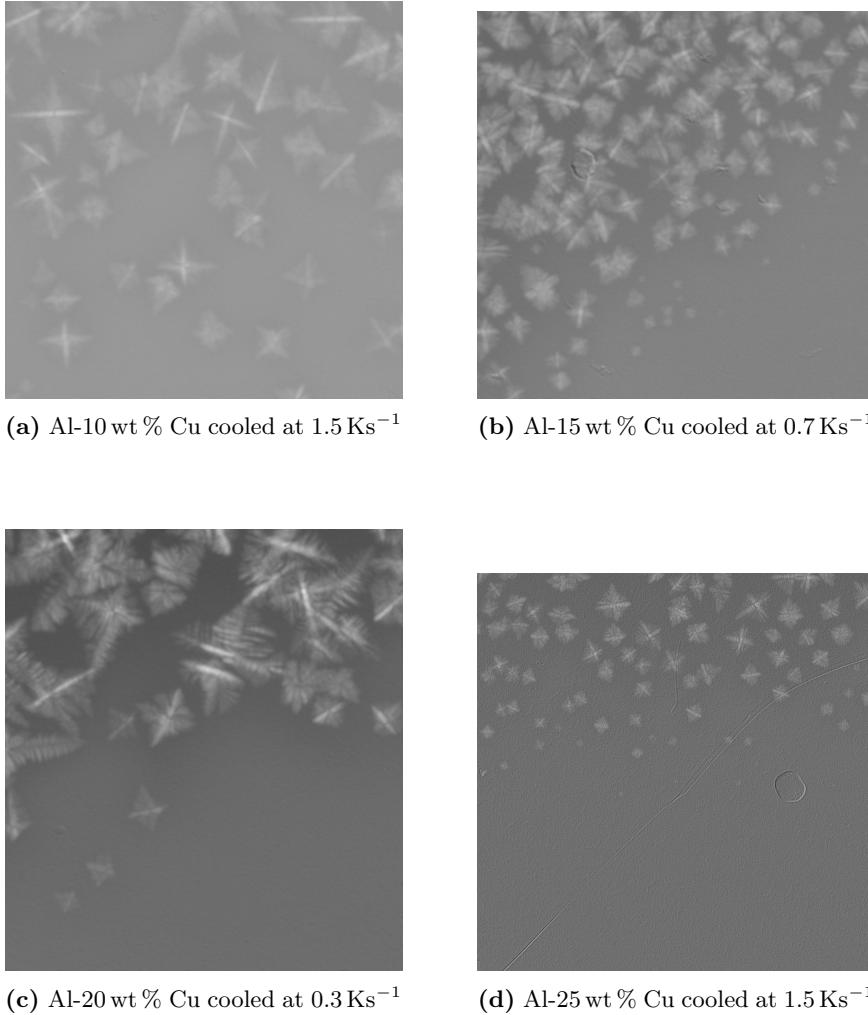


Figure 3.3: Selection of frames from different experiments within the dataset, showing the equiaxed grains approximately half way through the sequence.

To train a object detection model, annotations in the form of bounding box coordinates around the grains in the image are required. Annotations for 123 experiments were produced by the pre-deep learning grain detection model developed by Liotti *et al.* and accessed via a custom MATLAB class designed to store, handle, and retrieve data collected in X-ray radiography experiments. New scripts were written in order to extract the images to a JPEG format, and bounding box coordinates to a CSV, resulting in 13,887 annotated frames. Scripts were produced in Python to convert the CSV data into a COCO² style JSON annotation file, which will allow for annotation sets to be compatible with a variety of open source tools in the object detection field. In order to verify the bounding boxes,

²Further information on the COCO dataset is given in section 3.3.2.

3.2 Dataset

the VGG Image Annotator (VIA) was used, which allows for images and their associated bounding boxes to be visually inspected or edited [79]. During this exercise, several experiments were identified to have low quality annotations and were filtered out of the dataset, resulting in 9,909 annotated images. It should be noted that despite using a filtered set of annotations, they do include a small number inaccurate detections and some missing detections. Due to the size of the dataset and constraints of the development timeline, it would be inefficient to manually correct all annotations. Instead, one expects the more resilient nature of CNN based approach to pragmatically learn from the larger number of correct annotations.

3.2.2 Data Split

When initially prototyping the model, all frames were randomly shuffled and split by 80 % and 20 % to training and validation sets, without a separate test set. This setup was intended to be used solely for validating the training process and dataset scripts, alongside debugging any early technical issues. However, during this process it became apparent that all frames from a unique experiment should be in the same set. This requirement is necessary as there is very little visual difference between two successive frames. This can lead to frames in the validation set being too similar to that of the training set, resulting in an overestimation of model performance during evaluation between epochs³.

To produce a more appropriate dataset, a algorithm was designed to take into consideration the varying number of frames per experiment and output combinations of experiment numbers which represented a train/validation/test set split which was closest to a desired input. The outputted combination was then manually verified to ensure even distribution of the 4 different alloy compositions and cooling rates between the sets.

³Further details on model evaluation available in section 3.3.2 and epochs in section 3.3.1

3.2 Dataset

Set	Frames	Percentage of total
Train	8979	89.7
Validation	701	7.0
Test	329	3.3

Table 3.1: Details on the first filtered and split dataset

In order to make fair comparisons between models developed, and assess the final model performance, the test set was manually corrected using the List Annotator (LISA) software [80]. The final distribution of images over each set is detailed in **table 3.1**.

3.2.3 Data Processing Pipeline

Once the dataset has been defined and structured, several steps can be taken to enhance or alter features of the dataset in order to improve performance and precision. One feature of the annotations generated by the prior model was its lack of ability to detect crystals that are on the edge of the frame, and hence only a proportion is visible. In order to introduce these features into the data, an 150 pixel crop was introduced to the entire dataset, with the associated transformation made to the bounding box data.

One can use similar transformations to the data to artificially increase the size of the dataset, which is known as the process of augmentation. Augmentation primarily consists of two main classes of approaches, pixel-level transformations and geometric transformations. Pixel-level transformations include colour-based changes and noise introductions, while geometric transformations could include horizontal flips of the image and resizing the dimensions [81]. Augmentation has enabled datasets to be significantly increased, catering to deep neural networks which perform well on very large datasets [82], while demonstrating significant improvements to model precision [83–85].

To implement augmentations, a function was written using Albumentations, an open-

3.2 Dataset

source python library which allows quick implementations of augmentations to both images and the associated bounding boxes [86]. By iterating over different augmentation combinations and comparing performance, the following strategy used for training the final model: images from the training set randomly faced one of three options: a random crop to a 512×512 px. portion of the image, a resize of the original image to 1024×1024 px. dimensions, or a resize of the original image to 512×512 px. dimensions. 50% of the images would then have their brightness and contrast randomly altered between set bounds. This set of images was combined with the original un-augmented train set to double the training set size. Augmentations have been visualised in **figure 3.4**, noting that there is no visual difference in the 'Longest Max Resize 512' as the image is only resized to smaller dimensions.

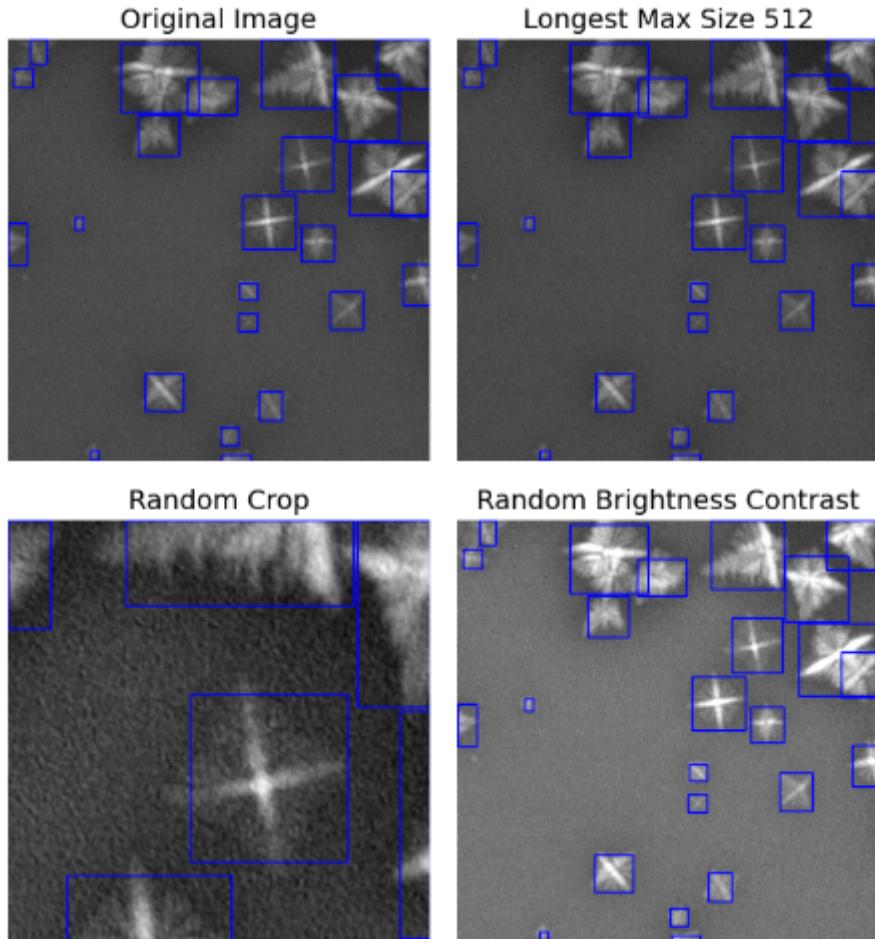


Figure 3.4: Visualisation of augmentations applied to an image from the dataset

3.3 Model Training and Evaluation

The final step is to convert the data in a form that can be processed and used by the Faster R-CNN model. For a neural network to be able to use the data, both images, bounding boxes, and image labels must be in the form of a tensor. A PyTorch specific dataset class is constructed which takes the data and converts each entry to a PyTorch tensor. The augmentation function is an argument of the dataset class, and images and bounding boxes are transformed as per the function. The class returns an image and an associated Python dictionary containing bounding boxes, the image label and the image area, all in a tensor format. With all the prior steps taken, the model can now be trained.

3.3 Model Training and Evaluation

Training a deep learning model is an iterative procedure, through which the training set is processed by the model to improve the accuracy of detections. This subsection outlines the theory behind training neural networks and explores how object detection models are evaluated both during the training and once the model is fully trained.

3.3.1 Training

In the training step of model development, the goal is to rectify the neural networks parameters in order to improve its ability to detect and locate grains within an image. For the Faster R-CNN network, this includes training both the RPN and the ResNet backbone. The model is initially loaded with pre-trained weights, based on the ImageNet dataset, an image classification and object detection dataset with 1000 object classes and 1,034,908 images labelled with bounding boxes [87]. Despite none of these images containing aluminium grains imaged through X-ray radiography, the model can make use of existing learnt features and reduce the training time, a technique known as transfer learning [88].

3.3 Model Training and Evaluation

During training, the model attempts to predict bounding boxes and classify each box. For Faster R-CNN, an equation denoted a multi-task loss function considers the both the performance of bounding box prediction and classification and returns a scalar value [66]. This is shown by the two terms in **equation 3.1**, where cls is classification and reg is bounding box regression. The loss function penalises the model for wrong classifications and inaccurate bounding boxes through adjusting the the parameters (weights) of the neural network in a process called backpropagation. During backpropagation, the gradient of the loss function is used to inform the model which direction parameters should be adjusted to decrease the scalar loss value, known as gradient descent [22].

$$\mathcal{L}(p_i, t_i) = \frac{1}{N_{cls}} \sum_i L_{cls}(p_i, p_i^*) + \lambda \frac{1}{N_{reg}} \sum_i p_i^* L_{reg}(t_i, t_i^*) \quad (3.1)$$

Once this has been completed for the entire training set, an epoch is complete, where an epoch is one step of training through the entire training set. Between epochs, the model then performs predictions on the validation set and evaluates the performance, outputting several metrics. The training process continues for several epochs. During the training of the aluminium grain detection model, 20 epochs resulted in these metrics converging to stable values. Training was conducted on an NVIDIA RTX A6000 GPU and took approximately 1 h 40 min for 20 epochs.

3.3.2 Evaluating Object Detection Models

Evaluating the object detection model is critical to monitor and demonstrate the ability for the model to identify and locate grains in images. It takes place both during training, specifically between epochs, and is used at the end of training to understand the final model performance. Before exploring the methods in which model evaluation is carried out, several key concepts should be defined in the context of the model's application.

3.3 Model Training and Evaluation

Precision is the proportion of correct grain identifications (true positives) compared to all the predictions made by the model. It defines the accuracy of the model when detecting grains. Recall is the proportion of correct grain identifications compared to the actual number of grains present. It measures the model's ability to detect all grains in an image.

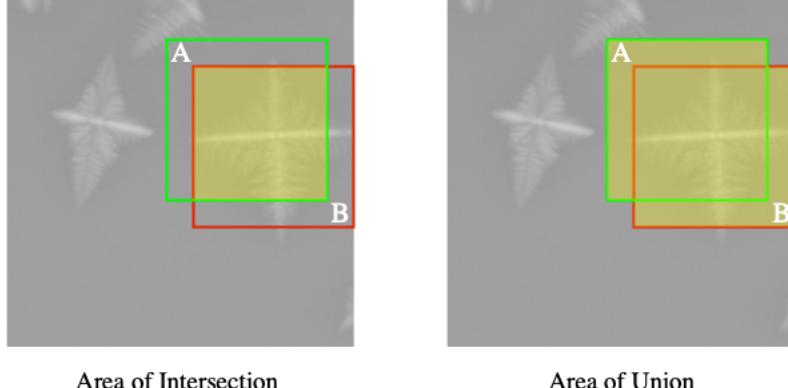


Figure 3.5: Intersection and Union for an example prediction (green) and original label (red).

In both these definitions, detections have been referred to as correct or incorrect, yet for an object detection task, it is important to explore how this distinction is made. Since detections are outputted as bounding boxes in coordinate form, it unrealistic to expect the model to predict the exact same coordinates as the annotations. To determine if a prediction is correct or not, the intersection over union (IoU) is used, defined in **equation 3.2** and visualised in **figure 3.5**. A IoU threshold value can be set to then determine if the detection is correct, and this value can be changed to evaluate precision and recall at different levels.

$$\text{IoU} = \frac{|A \cap B|}{|A \cup B|} \quad (3.2)$$

These three concepts are brought together in the form of a precision recall (PR) curve, which will be used as a key tool to visualise model performance. A precision recall curve plots precision on the y-axis and recall on the x-axis. Data points for the PR curve are gathered by generating predictions on a test set, where each prediction comes with

3.3 Model Training and Evaluation

a confidence rating between 0 and 1. A threshold is set for confidence, and a value for precision and recall for the entire set is found. This is repeated for various confidence thresholds, where a lower threshold would generally lead to an increased recall (less grains missed) but at the cost of reduced precision (more incorrect predictions) and a higher threshold would lead to higher precision and lower recall. For an ideal model, precision would remain as close to 1 as possible at all values of recall. The area under the PR curve represents a single value called average precision, and when it is averaged over all classes, is called the mean average precision (mAP).

The remaining metrics used all are defined through the Common Objects in Context (COCO) Dataset, and the related challenge metrics [89]. Similar to the ImageNet dataset discussed in **section 3.3.1**, the COCO Dataset is a large computer vision dataset with several associated machine learning tasks that can be trained from the set, one of which is object detection. The dataset can be used to compare and benchmark the performance of various models as they are developed. To ensure standardisation, COCO comes with 12 metrics, alongside a specific file structure for annotations and predictions. Of the 12 metrics, 6 are average recall (AR) based, and 6 are average precision (AP) based. Within these two categories, the metrics vary by looking at ranges of object sizes, counts per image and IoU ranges.

As previously mentioned, evaluation took place at two points in the development pathway. During training, evaluation on the validation set was implemented via a relevant reference script provided by torchvision. The script provided a CocoEvaluator class which is called during evaluation to produce the metrics [78]. In order to monitor these metrics and visualise their change over epochs, scripts were edited to collect and store the values. TensorBoard is a visualisation toolkit initially developed for TensorFlow [90], another large machine learning library, which was introduced to PyTorch and has been used in the

3.3 Model Training and Evaluation

model development to visualise and track converged metrics over many successive models.

After training was complete, models were evaluated using a software package called Object Detection Metrics written in Python, developed by Padilla *et al.* (2021) [91]. It takes test set annotations, and then model predictions, formatted in a COCO style JSON, and produces a variety of metrics, including the 12 COCO AP and AR values, alongside a PR curve. Since the software was written in Python, some modifications were easily made to manually export the PR curve data to a CSV so multiple curves could be plotted on the same axis.

The set of COCO metrics is just one example of many that can be calculated from detection data. They mostly arise from similar public datasets and their associated model challenges [92], one example being the PASCAL Visual Object Challenge which ran from 2005 till 2012 [93]. The PASCAL VOC mAP is the mAP calculated at an IoU threshold of 0.50, and has been used as one of the primary metrics for comparing trained models through the development of the grain detection model.

3.3.3 Measuring Performance Changes During Solidification Sequences

Once initial models were trained, more information was needed to understand at what stage in the reaction model performance would decrease, and the associated frames at these points. By doing so, it would give an indication on where the model can be improved further. Since solidification was occurring at different cooling rates, the time or frame number could not be used to compare reaction progress between experiments, therefore a different approach is required.

Using the test set annotations, which were considered to be the ground truth, the total area coverage by aluminium grains were calculated by finding the area of the bounding boxes for each frame. Percentage coverage was found by dividing this value by the total

3.4 Model Results

image area. Frames in the test set were then separated into different 5% interval groups, with the following distribution plotted in **figure 3.6**. By doing so, each frame in the test set now has an associated descriptor indicating the reaction progress.

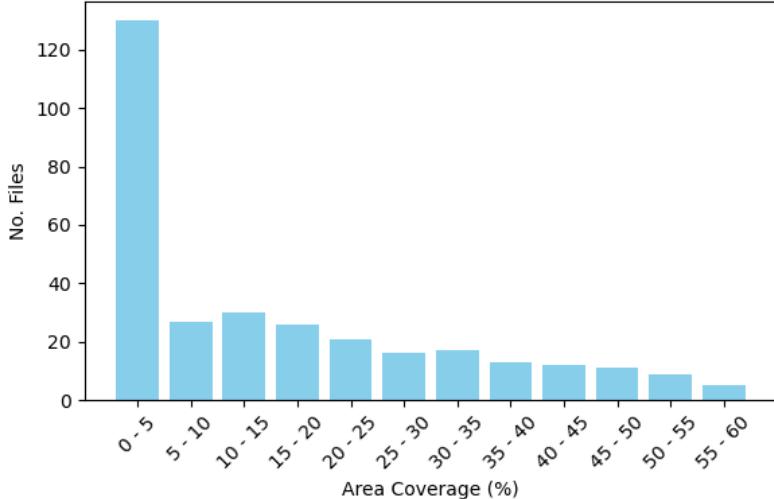


Figure 3.6: Distribution of frames by area coverage proportion intervals

Once the frames were isolated into different folders, a script was written to search for the relevant annotations from the CSV and make individual COCO JSON annotations for each coverage interval. The model made predictions for each interval and then the Object Detection Metrics tool [91] was used to produce both mAP values and data points for PR curves.

3.4 Model Results

To conclude this section, the final performance of the optimised and trained models will be briefly outlined and compared. The final project resulted in two models, one which aimed to perform well over all alloy types, and one specifically trained to perform well for 25 wt % Cu alloys. An overview of how the model has performed over the sequence will be presented to highlight the strength and weaknesses in the context of grain detection.

3.4.1 Reducing the Dataset for 25 wt % Cu Alloys

During model development, it was noted that alloys with a composition of 25 wt % Cu had the highest quality annotations produced by the previous detection model in 2018. This could be attributed to the visual features of the crystals having a more consistent uniform shape, alongside minimal overlap during the solidification sequence. The primary analysis of nucleation events which will be covered in this thesis all are also of Al 25 wt % Cu alloys, but applied to solidification conditions which the previous model struggled to detect well.

Model	No. Annotations
A	299,049
B	86,990

Table 3.2: Total number of annotations in training sets of respective models

A reduced training and validation set was created which used those alloys to investigate the effect of limiting the training data to higher quality annotations, or just to produce a limited but more accurate model specifically for 25 wt % Cu alloys. The test set remained the same to make a fair comparison between models. For any results detailed, the model with the entire dataset will be referred to as Model A, while the reduced set model will be referred to as Model B. The total number of annotations for each set is given in **table 3.2.**

3.4.2 Key Metrics

For the final models trained, a sample of metrics have been recorded in **table 3.3**. We can observe a higher mAP when using the PASCAL metric (IoU threshold of 0.5) for Model B but the COCO AP is better when using the larger dataset of model A.

Metric	Model A	Model B
mAP (PASCAL)	0.676	0.699
AP (COCO)	0.365	0.354
AP 75	0.365	0.332
AR 100 (COCO)	0.451	0.432
AR Medium	0.495	0.476

Table 3.3: Key metrics from trained models.

3.4.3 Precision-Recall Curves

Figure 3.7 shows the PR curves for both Model A and Model B. The curve generally suggests that model B maintains a marginally higher precision at higher recall values

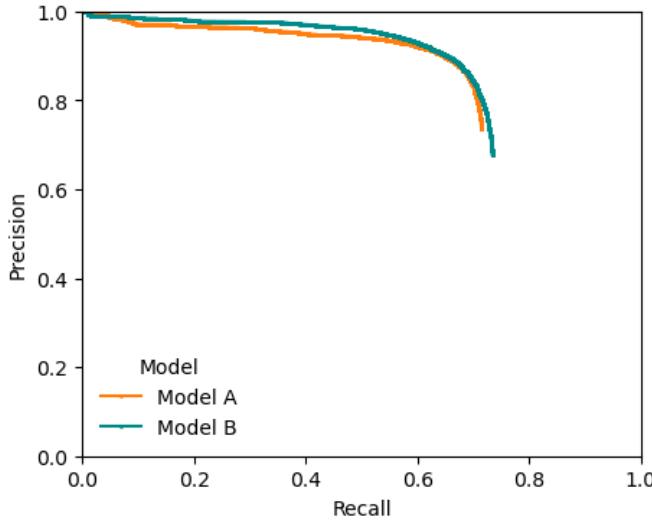


Figure 3.7: PR Curves for the two final trained models

Figure 3.8a shows a series of PR curves for Model B at different stages of the solidification sequence while **figure 3.8b** plots the PASCAL mAP of Model B for the different stages of the solidification sequence.

3.4 Model Results

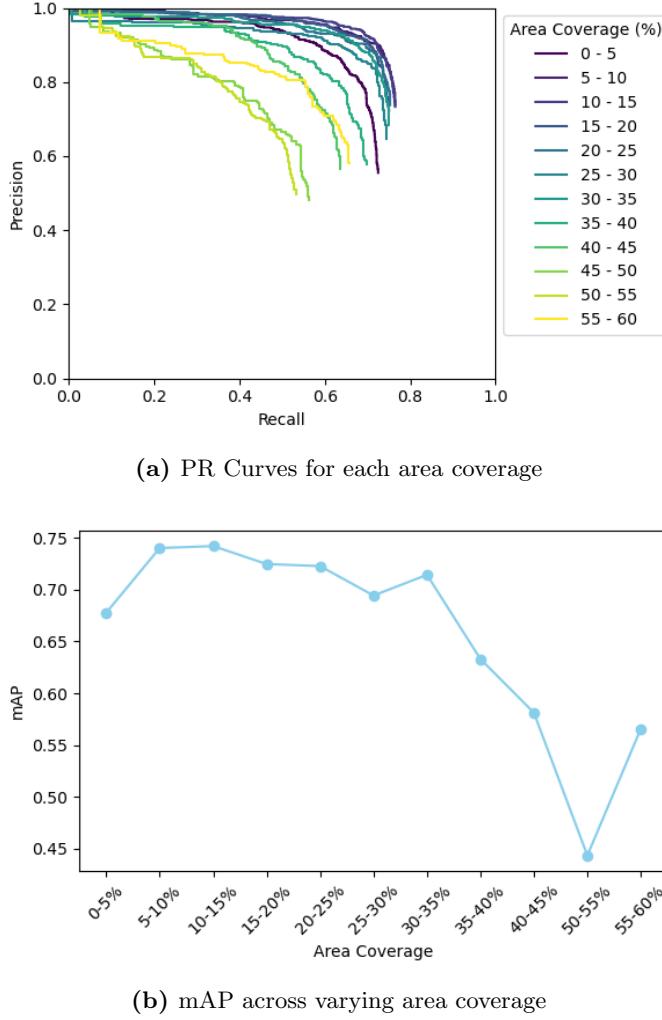


Figure 3.8: Investigation into Model B performance at different stages of solidification reactions

3.4.4 Understanding Model Predictions

For an radiographic sequence of 1000 frames, the model completed inference in 33.6 seconds, giving an approximate processing rate of 30 frames per second. This was conducted using an RTX A6000 GPU and Intel Xeon w9-3495 CPU, however inference speed will vary depending on compute available. This is contrasted to the previous model developed which took over 2 hours for a similar length sequence [13].

In **figure 3.9**, 9 images from a test set experiment have been included, showing both the evolution of the reaction and the model’s predicted bounding box for each crystal. A green box represents true positive prediction with an IoU > 0.5 for the associated ground truth annotations. It is noted that some boxes show appropriate identification of a crystal,

3.4 Model Results

but remain red as the IoU threshold has not been met. This discrepancy demonstrates an element of uncertainty in the human corrected annotations, as in some cases the model has made a better representation of the crystal's bounding box than the manually corrected annotations.

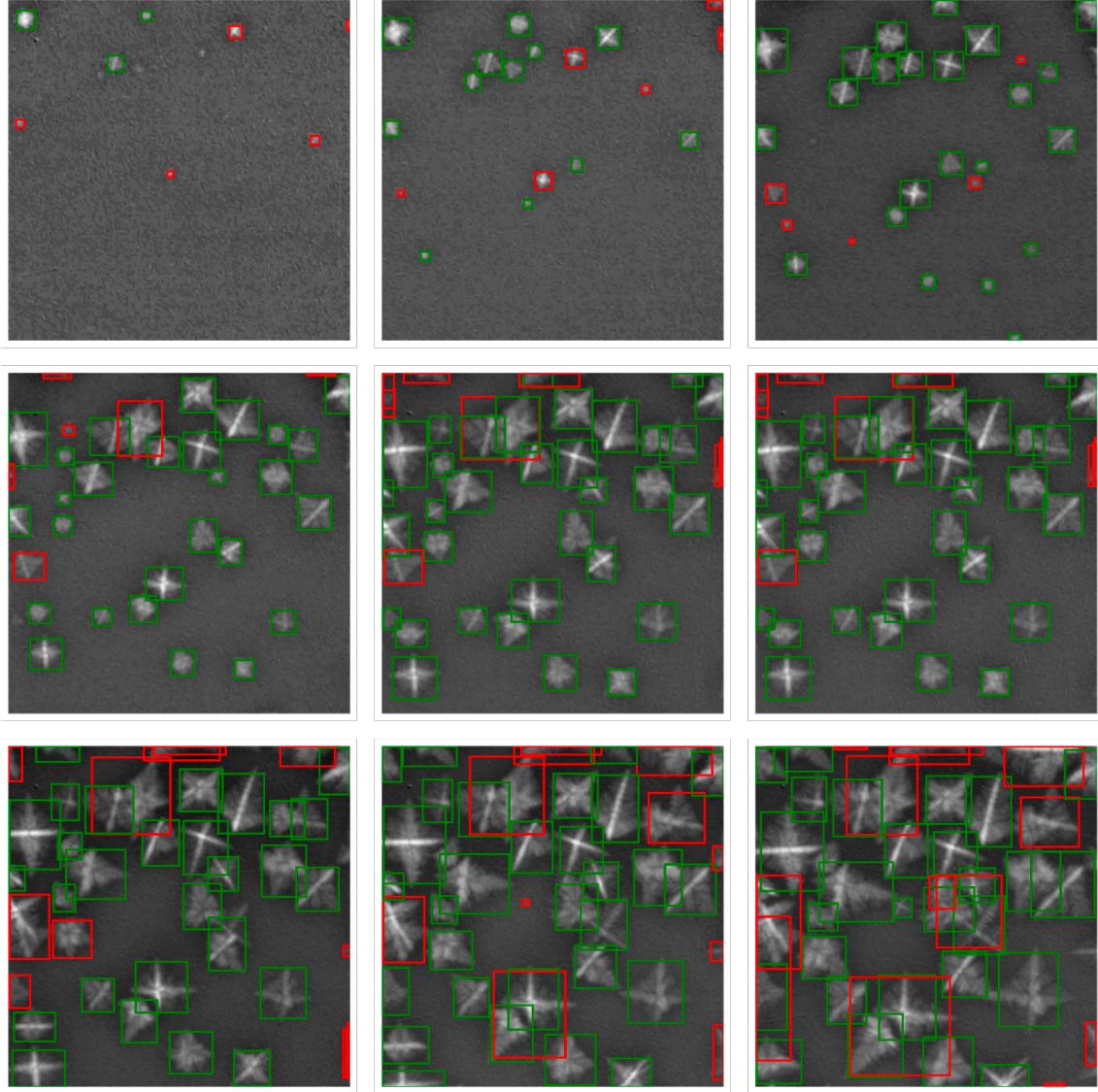


Figure 3.9: Model predictions over a solidification sequence with box colours determined by an IoU threshold of 0.5

Towards the end of the sequence, it is observed that closely overlapping grains leads to decreased predictive accuracy. This is seen through two mechanisms, the proposal of boxes which begin to cover two or more grains, and the drop out of individual detections when grains begin to overlap. This effect is observed across almost all sequences where overlap occurs, and can be attributed to several reasons. Firstly, the RPN is designed to

3.4 Model Results

make single proposals on distinct objects within an image, hence techniques exist within the model to prevent making multiple predictions on the same object. When overlap occurs, these techniques used designed to prevent redundant bounding boxes could lead to the suppression of true grains. Secondly, within the training dataset used, there are limited examples of correctly annotated overlapping grains leading to the model poorly generalising in these situations. Lastly, identifying individual grains from overlapping and even just adjacent examples results in a complex visual pattern. During the manual annotation of the training set, without prior context, identifying unique crystals during the later stages was a difficult task. It is plausible to make this comparison towards the feature extraction which the CNN is attempting to achieve.

This effect is quantitatively shown in **figure 3.8b**, where area coverage greater than 30% is where we begin to observe more overlap in the various sequences. As a result, the mAP is observed to decrease steadily. A discussion into how this could be mitigated in future iterations of the model is presented in **section 5.1.2**.

Chapter 4

Nucleation Event Analysis

4.1 Experimental Methods

The solidification sequences provided for analysis were conducted through the following method. Four samples of Al-Cu alloys with compositions of 10, 15, 20 and 25 wt % Cu with 0.1 wt % Ti for grain refinement were prepared by melting high purity Al, Cu and Al-5Ti-1B (wt %) using an induction furnace. The liquid metal was cast into cylindrical copper molds and cooled. Cylinders were sectioned into 0.3 mm discs and thinned through polishing down to approximately $200\text{ }\mu\text{m}$ thick. The foil samples were sprayed with boron nitrite to prevent oxidation and placed between two $100\text{ }\mu\text{m}$ quartz plates. The final sample compositions was inspected using an scanning electron microscope with energy dispersive spectroscopy to verify the chemical composition [13].

The enclosed samples were then placed into the Bridgman furnace solidification rig, which was designed to allow the X-ray beam to pass through. Thermocouples were placed inside the heating plates and measured the temperature to $\pm 0.1\text{K}$ accuracy. Radiography was performed at the ESRF using the ID19 beamline. A white-beam was generated using a single-harmonic undulator source, resulting in X-rays with a peak energy of 19 KeV. As

4.2 Analytical Methods

mentioned previously in **section 3.2.1**, the transmitted X-rays through the sample were collected using a GGG:Eu scintillator, where generated visible light was magnified and projected to a pco.DIMAX camera, and captured at rate of 6.67Hz [13].

These samples were completely melted and then cooled with different cooling rates and EMF stirring regimes, with several repeated experiments, resulting in 268 total sequences collected.

4.2 Analytical Methods

In order to maximise the range of cooling rates covered out of the data available, specific experiments were selected from sequences of Al-25 wt % Cu. In total, 19 radiographic sequences solidified with and without EMF stirring at cooling rates between 0.1 and 1.5 Ks^{-1} were analysed.

4.2.1 Grain Nucleation Counting

Raw images were pre-processed according to procedures outlined by Liotti *et al.* in 2018 and exported as in a JPEG format via. a bespoke MATLAB package [13]. A new python script was prepared to use a trained model to perform predictions on the images and save the bounding box coordinates and associated confidence scores to a CSV. When using this data for either quantitative analysis or visualisation, detections with confidence scores 0.8 were used.

For each experiment of interest, a the total number of detected bounding boxes (representing the number of grains) was plotted against time. **Figure 4.1** provides two examples of the raw value of box detections plotted against the time from the first grain nucleation event. These figures provide context for the level of noise present in the raw data produced by the model this dataset.

4.2 Analytical Methods

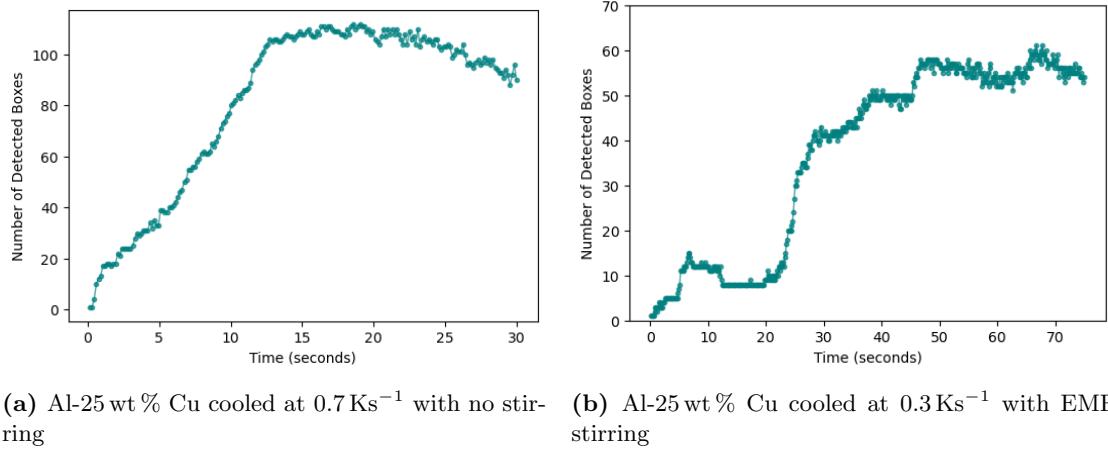


Figure 4.1: Two example experiment counts obtained from model predictions

4.2.2 Smoothing and Refining Grain Count Curves

The data series generated by the model were subject to several key processing steps in order to remove noise and represent the number of grains nucleating in the system. The steps implemented were inspired by the processing steps taken by Liotti *et al.* in 2018 [13].

A monotonicity correction was applied so that the grain number only increased when a new grain was detected, but did not increase when a grain moved out of the field of view or due to detection drop out by the model. Detection dropout typically would occur for 1 or 2 frames before re-detection, while crystal movement out of frame represents an unavoidable feature of the dataset. For stirred experiments with low cooling rates (0.1 and 0.3 Ks^{-1}), the monotonicity correction was not applied due to periodic movement of grains in and out of the field of view, which is detailed further in section 4.4.2.

The series were then filtered with a moving average and fit with a spline function. The moving average was used to smooth out localised jumps in crystals detected after detection dropout had occurred before, while the spline was employed to convert the data into differentiable function. The derivative was calculated and plotted as the crystal formation rate. The primary goal of the formation rate was for the detection of nucleation bursts, and hence values below 0 were omitted for visual clarity.

4.3 Bursts of Nucleation

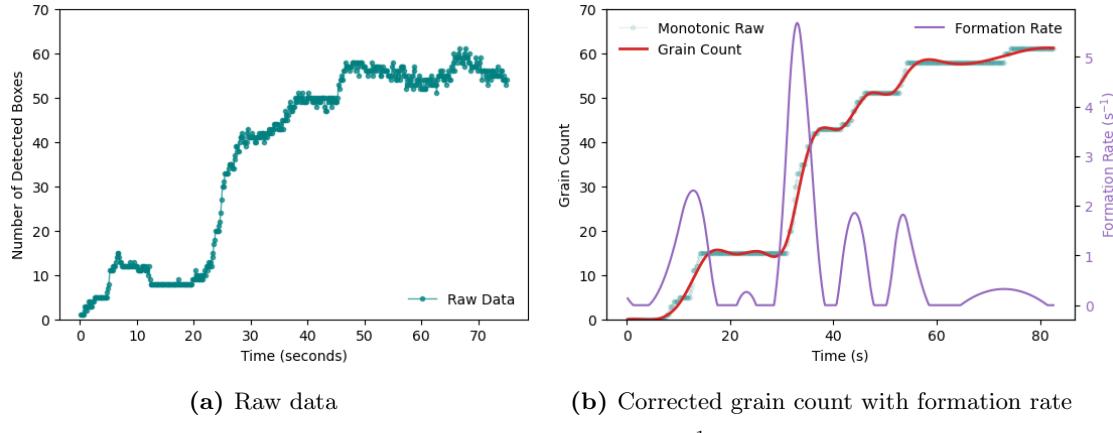


Figure 4.2: Al-25wt% Cu cooled at 0.3 Ks^{-1} with EMF stirring

Figure 4.2 shows both raw box count data and the associated smoothed representation of grain count as a result of the processing steps detailed above.

4.3 Bursts of Nucleation

Figure 4.3 shows the nucleation behaviour of two solidification experiments of Al-25 wt % Cu at a cooling rate of 0.3 Ks^{-1} without EMF stirring. Four nucleation bursts are observed both on the grain count curves and the formation rate curve. On the formation rate curve, certain bursts have been noted with a star, which is due to a burst occurring at the edge of a frame. This results in the rate increase being not as distinct as the nucleated crystals move into the frame from buoyant flow rather than nucleating within the field of view.

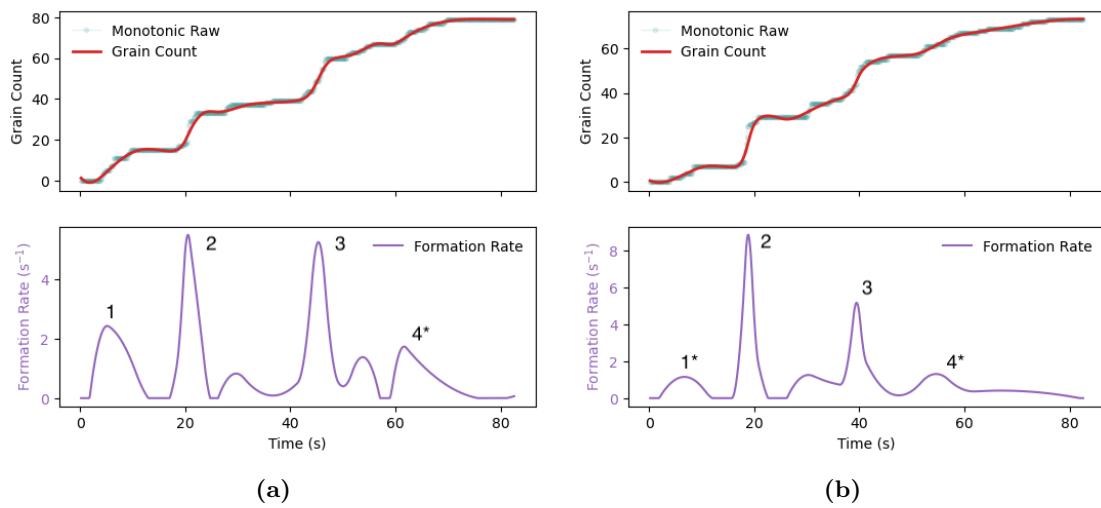


Figure 4.3: Al-25 wt % Cu cooled at 0.3 Ks^{-1} with no stirring

4.4 EMF Stirring

The observed bursts match what was expected and presented by Liotti *et al.* in 2018, validating the models ability to quantify rate changes in nucleation.

4.4 EMF Stirring

4.4.1 Comparing Burst Behaviour

Having demonstrated nucleation bursts at 0.3 Ks^{-1} , three experiments were obtained at the same cooling rate which were subject to various stirring through EMF. Their grain count curves and formation rates are shown in **figure 4.4**. In **figure 4.4b & 4.4c**, we see that the nucleation bursts are no longer present and by visual inspection, the average nucleation rate is faster. In **figure 4.3a**, bursts are observed.

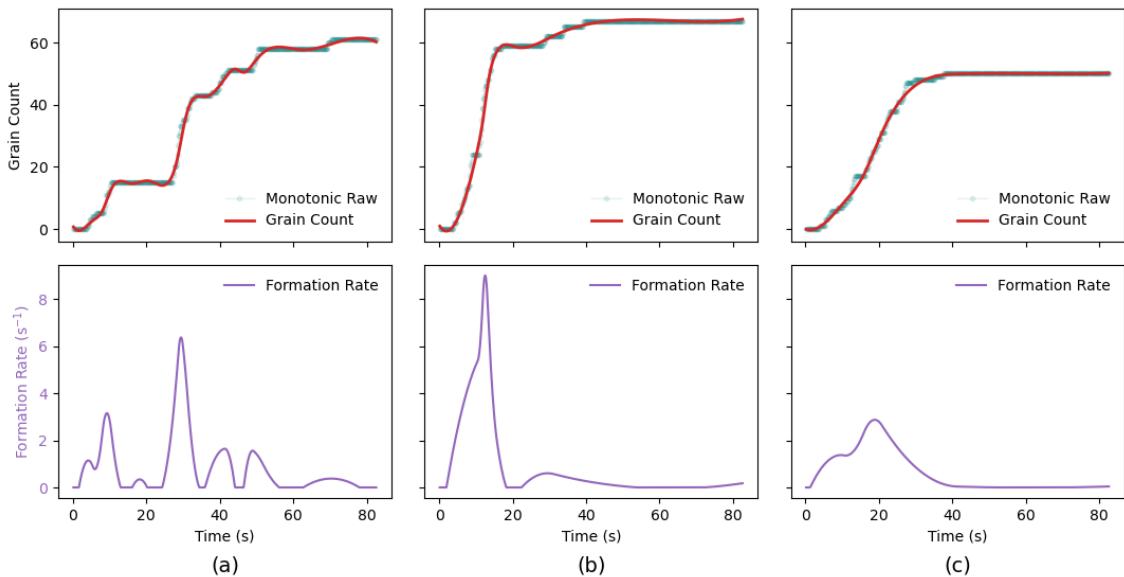


Figure 4.4: Al-25 wt % Cu cooled at 0.3 Ks^{-1} with differing stirring regimes

When considering the different EMF stirring conditions, **4.4a** and **4.4b** were both subject to a field using a sine waveform, however, **4.4a** has the EMF force applied in the opposite direction than **4.4b**. The experiment in **4.4c** is stirred with an exponential base shape waveform with manual modification of the amplitude, representing a more unstructured and noisy stirring method.

4.4 EMF Stirring

When examining the videos of the solidification sequences, both **4.4b** and **4.4c** show the majority of grains flowing upwards and oscillating horizontally at the initial stages of grain growth. After a critical size is reached, the buoyant upwards flow stops and the grains grow in a fixed position. With **4.4a**, a smaller proportion of the grains move follow this pattern of upwards motion and horizontal oscillation, with more reaching their stationary position earlier in the sequence.

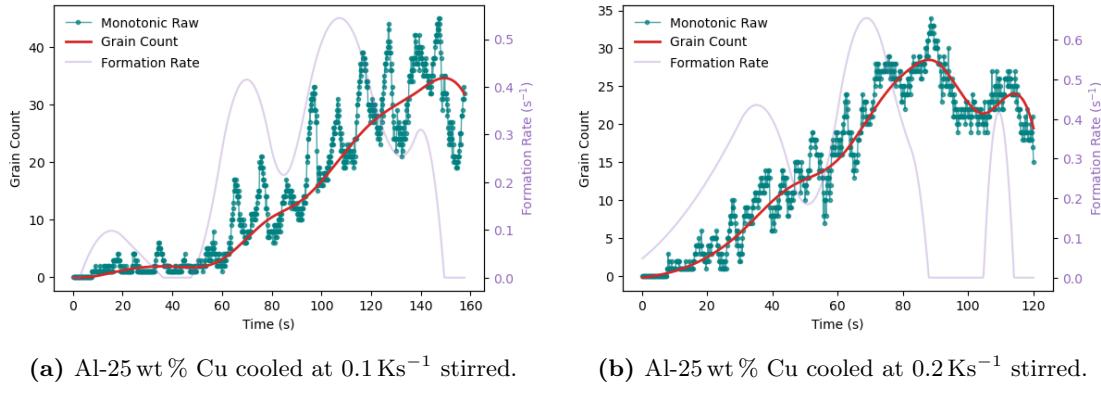
While inconsistent EMF stirring conditions represents an unavoidable limitation within the available data, the preliminary analysis conducted combined with qualitative observations suggests the following: where stirring is successfully applied to the majority of the crystals, the presence of nucleation bursts is seen to reduce.

One could hypothesise that this may be due to several reasons: firstly, stirring promotes a more homogeneous distribution of Cu where the SSNZ would be, leading to a reduction of undercooling required for nucleation to take place; secondly, an increased dispersion of inoculants such as Ti used in this system, which would lead to a more efficient nucleants being present adjacent to where nucleation has occurred; and finally, the forced convection of particles could help decrease localised temperature gradients introduced after nucleation occurs. In order to expand upon these conclusions, a larger set of solidification experiments would have to be conducted for each specific stirring regime used. By varying the magnitude of forced convection, the frequency of bursts could be used to indicate the effect of stirring on the SSNZ.

At this cooling rate and stirring regime, the model performs very well, clearly detecting most crystals in all the reaction sequences.

4.4.2 Low Cooling Rates

Low cooling rates refers to experiments conducted at 0.1 Ks^{-1} and 0.2 Ks^{-1} within the data available. In these experiments, stirring is conducted through a field controlled by a sine wave, with a $\sim 3\times$ greater amplitude and reduced frequency than experiments conducted at 0.3 Ks^{-1} in **section 4.4.1**. The greater amplitude results in crystals moving completely in and out of the field of view, which is shown by the distinct character of the raw count shown in **figure 4.5**. The monotonic correction was not applied, and instead the moving average was applied with a window greater than the period of oscillations in and out of the frame, alongside an increased smoothing factor to the spline interpolation. When visually verifying the detections, the smaller crystals are captured well by the model when in frame, however in the later stages, when grains stop moving, detections begin to permanently drop out due to overlap and close proximity, as previously discussed in **section 3.4.4**.



(a) Al-25 wt % Cu cooled at 0.1 Ks^{-1} stirred. **(b)** Al-25 wt % Cu cooled at 0.2 Ks^{-1} stirred.

Figure 4.5: Grain count curves and nucleation rates conducted at 0.1 Ks^{-1} and 0.2 Ks^{-1}

To conclude, the model is able to capture the result of high amounts of EMF stirring, however the grain count curves are mostly representing the crystals which are still in motion, rather than the total number of grains. This could be attributed to both limitations in the data accessed, alongside previously acknowledged decrease in dropout when counting crystals in the later stages of nucleation.

4.4.3 High Cooling Rates

For 0.7 Ks^{-1} and 1 Ks^{-1} , two experiments for each cooling rate were accessed and had grain counts produced. The experiments at 0.7 Ks^{-1} were stirred with the same regime to those at 0.2 Ks^{-1} and yielded good model performance during the early stages of nucleation. Since the cooling rate is faster, the majority of crystals nucleate in these early stages where grains show convective motion, and we can consider the grain count curves to be accurate representations of nucleation rates. It should still be noted that the model starts to lose detections as grains become larger and stop moving.

For the sequences cooled at 1 Ks^{-1} , a pulsed EMF stirring regime was used. Grains in these experiments are initially detected well by the model, however the forced convection leads to many leaving the field of view. This led to the grain count curve potentially representing a lower value than what has been physically nucleated. When considering the visualisation of the solidification, it is noted that in one of the experiments, the majority of crystals have been nucleated prior to the first EMF pulse. Between the two experiments, the maximum number of grains detected varies by 27.9% which is contrasted by sub 10% count differences found in experiments conducted at 0.3 and 0.7 Ks^{-1} .

4.4.4 Average Nucleation Rates

In the previous sections, each cooling rate available has been presented, while considering the various EMF stirring conditions used, distinct features of the experiments, and the ability of the model to appropriately quantify the nucleation behaviour. To summarise, cooling rates of 0.1 , 0.2 , and 1.0 Ks^{-1} produced grain count curves which support the qualitative observations which could be made, but lacks the level of detail necessary for any theoretical assertions. Contrary to this, cooling rates of 0.3 and 0.7 Ks^{-1} have stronger quantitative representations of the nucleation behaviour.

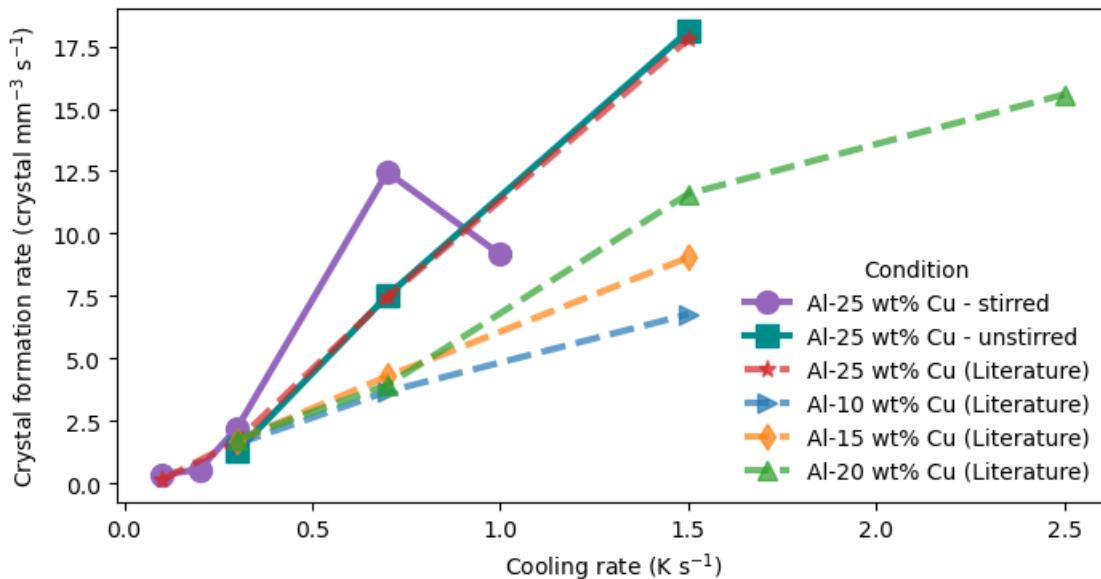


Figure 4.6: Average crystal formation rate for a variety of literature values [13] and the analysed series for Al-25 wt % Cu unstirred and EMF stirred

By considering the maximum grain count, the time taken to reach said point and the volume of each sample, average nucleation rates can be found for each experiment. **Figure 4.6** plots the literature values for a range of Al-Cu alloys, alongside two new series of data produced using the new model. The series produced for Al-25 wt % Cu matches the literature values produced by the pre-deep learning model developed by Liotti et al. (2018), providing further validity for the model on specific experiments. When considering the stirred series, the data points presented at 0.3 and 0.7 Ks^{-1} provides a preliminary result of increased average nucleation rates due to the presence of EMF stirring in the melt.

The reasons behind this effect are likely to be related to those which led to the removal of nucleation in bursts. A more homogeneous solute distribution, increased dispersion of nucleates, and reduction of thermal gradients all lead to the reduction or elimination of the SSNZ by broadly enabling nucleation to occur more easily. Therefore, it is logical that the average nucleation rate would increase, a trend which has been observed by the preliminary results.

Chapter 5

Future Work

5.1 Improving The Grain Detection Model

The work presented in this thesis has demonstrated greater generalisation in a model's ability to detect grains, and while doing so, demonstrated efficiency advantages provided by deep learning approaches. However, this represents a step towards a wider goal for a tool which can be used to perform analysis over a wide range of experimental conditions, and assist with a range of research goals within solidification. This includes different X-ray radiography set-ups, including lab based X-ray sources with significantly lower quality images, but higher availability, alongside using similar models for different alloy systems. The following sections outline the next steps which have been realised from the development of the existing model under the time constraints of the project, and should assist with the future steps taken towards the end goal.

5.1.1 Training Dataset

The training data used to train the model was produced by the previous machine learning model developed by Liotti et al. (2018) [13] without human intervention for correction.

5.1 Improving The Grain Detection Model

This was done to prioritise the amount of annotated data over perfect annotations, alongside limitations in the amount of time available to manually correct a dataset of over 8000 images. The new model presented in this thesis can be used in a similar style to generate annotations which may be less time intensive to correct. Another approach could include utilising a proportion of images from each sequence, for example every third image, leveraging the feature that each successive image is visually very similar. This would effectively decrease the dataset by a third and the required time for the associated annotation task. Higher quality annotations should help improve both recall and accuracy of detections, especially where the model has been observed to fail with more complex features in the dataset.

5.1.2 Overlapping Grains and Late Sequence Detection

In order to improve detection between adjacent and overlapping grains, two main approaches are considered. Firstly, increasing the number of annotated frames in the later stages of reactions. Example annotated frames in the training dataset were scarce, due to similar difficulties in the previous model to detect during these stages. This would have to be performed manually. However, as discussed in **section 3.4.4**, the visual features at this stage become less distinct, representing a problem for both humans and a CNN. With an increase in training data on overlapping grains, there is no guarantee the model would be able to make predictions to a usable level.

Consider the largest red box in the left side of **figure 5.1**, the model has correctly identified one of the grains, but how many more exist within the larger box? On the right hand side, the several earlier frames are included, and it becomes clearer that 3 total grains exist within that region. Just as context guides humans to understand more than what is immediately available in an image, techniques such as object tracking can be used to provide context to predictions. Information about the previous positions of bounding

5.1 Improving The Grain Detection Model

boxes are used in conjunction with the proposals made by an object detection model, helping prevent drop out when the detection model fails to identify grains in a sequence. A preliminary attempt was performed to integrate the DeepSORT tracking algorithm [94] and showed good results with limited experiments, hence further work is required for a robust implementation.

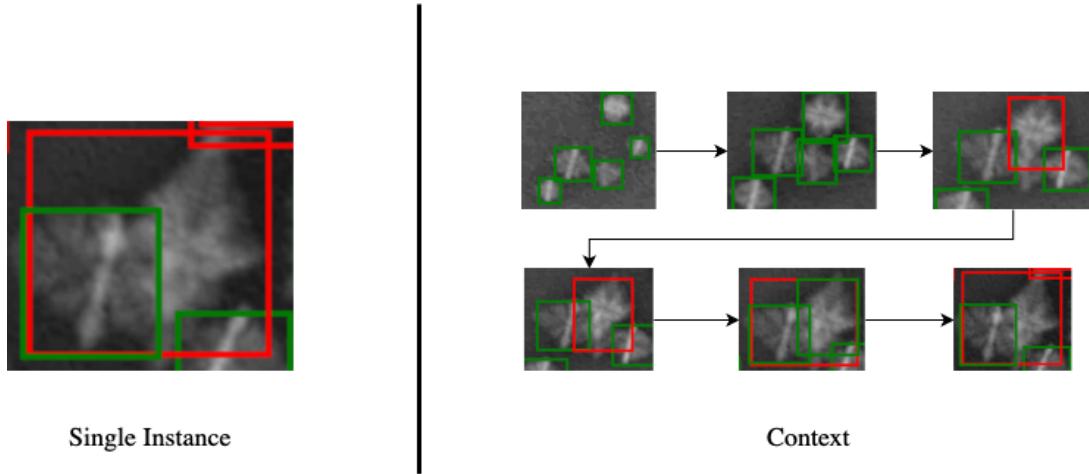


Figure 5.1: An example of where context helps one understand the number of grains present

Object tracking would also result in unique identification for each grain during a sequence. This would enable data to be gathered for the position of each grain during the sequence, a feature which would provide value when considering further work into investigating convection.

5.1.3 Grain Segmentation

Segmentation would result in more complex identification of each object within a frame, predicting the exact boundaries of a grain. An example of segmentation is shown in **figure 5.2**. Model architectures such as Mask R-CNN provide a clear route towards implementation [95], however require additional training data in the form of more complex polygons to define the boundaries of each object, called masks. A successful model would lead to relatively accurate data on the size of grains, which also would be valuable for further research within the field.

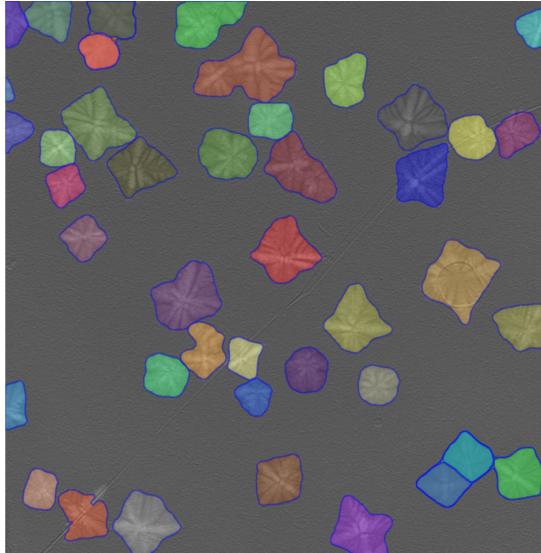


Figure 5.2: Example segmentation of a frame from the dataset, generated by the Segment Anything model [96]

5.1.4 Future Model Architectures

As the field of computer vision and deep learning continues to develop at a rapid pace, newer approaches to both object detection, tracking and segmentation has the potential to significantly change both the direction and pace of development for future work on this project. The introduction of transformer based models has demonstrated record benchmarks for object detection models. Successful implementation could assist with the goals of further generalisation, improved handling of sequential images, and integration of highly diverse datasets. An example is provided by **figure 5.2** generated by the Segment Anything model, a transformer based model presented in 2023 which has performed segmentation on a frame with only one noticeable mistake [96]. The model has been designed to perform one-shot segmentation, meaning it can be applied without specific training on a dataset. It should be noted that the success is partly due to the chosen image having visually distinct features and no overlap, however still demonstrates the potential if a similar model if it was to be fine-tuned on a full dataset of radiographic solidification sequences.

5.2 Investigating Convection

In order to present a more vigorous investigation into convective effects, future experiments should be conducted with set EMF stirring regimes, and repeated over multiple cooling rates per sample. With the preliminary results presented, the sinusoidal EMF applied to Al-25 wt % Cu cooled at 0.3 Ks^{-1} in **section 4.4.1** showed the most difference between its unstirred equivalent. Using this same stirring over more experiments and over the range of cooling rates would enable more reliable data points for average nucleation rate vs cooling rate plots.

It can also be noted that to ensure more reliable counts, stirring techniques should aim to keep grains within the field of view available during the experiment. If future work requires stirring of such magnitude, the field of view may need to be adjusted to accommodate the detection model.

The current understanding of what changes in EMF stirring methods may cause in nucleation rates is still preliminary, which means previously, experiments were conducted to maximise the range of conditions captured instead of repeating sequences. The significant decrease in inference time for the grain detection model means that early analysis can now be performed at the facility itself, informing decisions on experimental conditions during the allocated beam time.

Chapter 6

Conclusions

This thesis aimed to investigate convective flow in aluminium-copper systems during nucleation and solidification through the development and application of a deep learning object detection model. The Faster R-CNN architecture was employed and trained on a large dataset of annotated *in situ* synchrotron X-ray radiography solidification experiments, demonstrating improved generalisation over a range of experiments while reducing analysis time.

The model was used to perform analysis on 19 experiments, both with and without forced convection through EMF stirring. Nucleation occurring in bursts were verified in 2 experiments, producing average nucleation rates that agree closely with literature. Preliminary results were presented for solidification sequences with EMF stirring, quantitatively demonstrating the loss of nucleation burst behaviour and comparatively faster nucleation rates for the first time.

Despite these results, the developed model encountered difficulties in detection during later stages of a sequence, where overlapping grains resulted in inaccurate predictions. Future work towards object tracking and improving the training set should widen the applications

5.2 Investigating Convection

of the model, and enable more quantitative studies of solidification phenomena.

The investigation into convective flow was limited by sequences available, but now provides insight into which EMF stirring regimes and cooling rates should be used for future experiments. By doing so, more robust evidence can be presented on the effects of convective flow on nucleation rates.

To conclude, the work presented represents an iterative step towards the wider context of understanding how one can modify solidification conditions to improve efficiency processing metals. The deep learning model lays the groundwork and demonstrates the potential of using modern computer vision approaches to improve analytical efficiency for *in situ* X-ray radiography studies of solidification.

Chapter 7

Project Reflections

7.1 Ethics and Sustainability

As outlined in the engineering context, the thesis presented is working towards improving efficiency in metal processing with the specific context of recycling. Reducing the processing time and the steps required to take scrap metal to a desirable product would decrease the overall energy usage and resulting lower carbon emissions.

Additionally, potential advantages from better microstructural control can lead to an improvement in final material properties from recycled metals with lower overall energy consumption. This will aid the reduction in cost of high quality recycled metals, making it more desirable for industrial and commercial use, hence decreasing the demand for newly extracted metals. In the case of aluminium, recycling only requires 5% of the energy used compared to extraction from bauxite ore [97].

While this provides a long-term motivation for the sustainable prospects, it is important to think about the immediate impact produced by the work conducted in the project, particularly the carbon footprint. Given the scope of the project, estimating energy usage

7.1 Ethics and Sustainability

through all computational equipment used may not provide an accurate representation of total footprint. This is due to the large number of variables that are difficult to record without established monitoring from the commencement of the project. Instead the energy usage of GPUs will be presented, noted as a particular area of concern with the exponential growth of machine learning and artificial intelligence through very large models [98].

By looking at the total number of times the model was trained, 42 runs were conducted on the original workstation using a QUADRO RTX 6000 GPU with an average run time of 8 hours, and 39 runs on the new workstation using a RTX A6000 GPU with an average time of 2.5 hours. With the two GPUs using 250W and 300W respectively, we can estimate the total GPU energy usage to be 113 kWh. The average carbon intensity of the UK grid over 2023 was 163 gCO₂/kWh [99], giving the total carbon footprint of training to be 18.3 kg of CO₂. For further context, an average UK household uses 7.4 kWh of electricity (not including gas) [100], giving training energy usage to be roughly equivalent to 15 days of equivalent household electricity usage.

This value should be acknowledged as sizeable but a cost which could be considered necessary to help reduce significantly larger industrial emissions. Where possible, future work could look towards identifying renewable utility providers to further reduce emissions.

Ethical consideration of the work carried out is equally crucial. Two areas identified as relevant to this thesis are the potential for biases to form in deep learning models, and ensuring transparency and reproducibility of methods carried out. Bias in DL models is commonly explored in social sciences, however should be considered in this application also. There exists a risk that the model could be over-fitted to the data provided, producing inaccurate results on new datasets taken from different X-ray sources or detectors. To mitigate this, the current iteration of the model has been identified as appropriate for the current experiment series from which the training set was formed. Within this

experimental series, a test set and visual inspection was used to verify that the model was not biased towards a certain experiment or condition. In future applications, care should be taken to obtain a wide range of different experimental set ups in the training set before publishing the model for other users.

When considering transparency and reproducibility, the thesis represents the main body of text to explain and justify the approaches taken towards the model development. Key scripts and documentation will be published on GitHub without model weights for future members of the research group or wider community to use. Trained model weights and dataset will remain proprietary to the research group until suitable for wider release. Publishing relevant scripts to public repositories ensures transparency, allowing others to verify the steps taken and explored in this thesis.

7.2 Project Management

7.2.1 Defining a Project Scope and Structure

The prior work conducted by my supervisor, Dr. Liotti, meant that at the very start of the project there were several identified directions in which deep learning could be applied to the broad range of *in situ* radiography experiments conducted over the last 10 years. After discussions held with my co-supervisor, Prof. Zisserman, we identified that building upon the object detection model was a realistic task given the project timeline.

In terms of developing and refining the model, the task remained open-ended, meaning I aimed to iterate over the model as far as possible until the 2nd half of Trinity term. By then, the aims shifted towards using what had been produced to analyse the scientific case study of convective flow. We initially envisioned a trained Faster R-CNN model to be developed, and then if results were successful, attempt to use several more recent object detection models and investigate what yielded the best results and performance.

7.2 Project Management

In the event that the development of the model was unsuccessful or not realistic within the timeline, the project was planned to shift towards using the older model developed by my supervisor to conduct a study on previously un-analysed data in a different context.

Regarding the structure of guidance and communication, bi-weekly meetings were held with: Dr. Liotti, who was the main departmental supervisor; Prof. Zisserman, who guided the direction of the deep learning within the project; and Dr. Arteta, who offered technical support towards the model development. Work was presented at each meeting, and advice was given to form the main objectives to be completed for the next meeting. The method in which tasks were implemented and completed was conducted alone, allowing for complete control and responsibility over scripts and data produced.

A lab book was not explicitly used throughout this project, however digital note taking methods were used as a daily log to keep track of progress and make notes throughout the model development. This was more useful as it was essential that the large number of code snippets and data points could be quickly indexed for reference.

7.2.2 Personal Development

The project presented a unique opportunity to learn and develop skills within the field of machine learning through an extended project, something I had been previously considered pursuing outside of academic work. Having had prior experience coding in Python, I initially had confidence in my ability to handle the technical aspects of the project. However, understanding how to properly use a wide range of libraries and open-source tools in both machine learning and image data handling represented a learning curve.

This also represented the first time I was handling large datasets, both in images and annotations, where I quickly learnt the importance of clarity in version control and record keeping. Time-consuming mistakes were made, especially when manually editing test set

annotations and merging them with master annotation files containing all the bounding box information. This lead to certain tasks having to be repeated several times when proper version control could have ensured these tasks only needed to occur once. This also applied to code produced, where by the end of the project countless scripts and notebooks existed, used for different steps in model development, training, data handling, analysis to name a few. Each of these scripts had various iterations which meant a systematic approach need to be taken to ensure older versions were not mistakenly used or implemented in wider work.

The initial use of MATLAB to process and extract images from the existing database also represented some challenges as it was a programming language I was less familiar with. However, the plentiful resources online and support from supervisors allowed for the image retrieval to be completed.

Regarding time management, the structure of regular meetings each with defined objectives, meant it was easy to allocate sub-tasks both on a daily and weekly timeline. Towards the later stages of the project, many of the tasks involved training models with various parameter or dataset changes, for up to 8 hours at a time. In order to avoid waiting for these tasks to finish during the day, longer training scripts were prepared during the day and scheduled to run overnight. This allowed me to spend the day interpreting results and understanding what changes needed to be made so that they could be run the following night. Learning to adapt to a task where certain steps have time constraints beyond my control has been a valuable experience.

7.2.3 Project Timeline

During Michaelmas, the main focus of the project was to extract images to produce the dataset, and start to prototype scripts which would train the model. The dataset creation

7.2 Project Management

stage took much longer than anticipated due to certain graphical issues in MATLAB leading to images being exported with white margins. This was combined with the image extraction method itself being a relatively slow process, representing an unavoidable technical limitation. The allocated time of up to 3 weeks for this task ended up taking approximately 2 months to complete. During this time, I completed several tutorials on PyTorch and object detection in order to reduce time spent when the data was ready. Significant time was spent on setting up software and development environments for Python and PyTorch due to some incompatibilities with the Windows workstation I was initially using. Most issues encountered had alternative approaches available, but more time was allocated to implement these solutions.

By the start of Hilary term, the majority of the images were collected and prepared. Significant time was contributed towards the handling of annotation file structures. Errors during converting formats were persistent due to the presence of invalid bounding boxes in the data which were difficult to initially detect. The first attempts to produce a trained Faster R-CNN model were completed by the end of Hilary term. This was paired with understanding that a new workstation would be procured with significantly more powerful computational resources, including four modern GPUs and running a Linux operating system.

By the start of Trinity term, the new workstation was set up and immediately reduced training time from around 6 hours for the full dataset to 1.5 hours. This was incredibly useful as the majority of the work from here onwards focused on iterating over different augmentations, conditions and parameters that slowly improved the model accuracy. It was now possible to schedule several training scripts running on multiple GPUs during the day and overnight. At this point, I was advised to allocate most of my time to maximising the precision of the Faster R-CNN model, as this would be most beneficial to future work

conducted.

As the second half of Trinity approached, experiments were identified and collected for analysis. The analysis was performed with few issues and yielded promising results forming the basis of chapter 4 of this thesis.

Reflecting on process, the majority of issues arose from programming related bugs and errors, many of which taking significantly longer than expected to fix. I believe I could have taken a more proactive approach to seek advice from supervisors and their colleagues instead of spending days or weeks trying to fix them myself. Additionally, looking back at the unforeseen time spent on extracting data, a better approach could have been employed to anticipating future challenges when developing the model. With many public datasets available, I should have used the time to develop and train a model on different data, which I could have then quickly adapted to the relevant dataset once it became available.

7.2.4 Project Management Forms

PROJECT MANAGEMENT FORM 1

Part II Project Description Form

After discussion with your supervisor YOU should complete this form and send a copy to the Undergraduate Studies Office at undergraduate.studies@materials.ox.ac.uk by Friday of 0th week of Michaelmas Term.

Name: Rohit Abraham **College:** Corpus Christi College

Address for correspondence: Corpus Christi College, Merton Street, OX1 4JF

Contact telephone number: +44 7759 238190

Title of project: Artificial Intelligence for multi-modal X-ray Imaging of metal solidification.

Supervisor: Enzo Liotti

What are the objectives of the project in order of priority?

- Develop a deep learning model for the analysis of solidification in Al-Cu alloys.
- Use the model for a scientific case study analysing existing data.
- Further development of the AI model for more complex analysis.

List the major milestones that must be accomplished in order to meet the objectives of the project:

- Set up technology access for the development and training of the deep learning model.
- Initial training for PyTorch (self-learnt and supervision from Carlos Arteta)
- Access and Labelling of data for the training dataset.
- Development, training, and testing of the deep learning model.
- Selection of dataset for scientific application.
- Analysis of data and quantification of solidification phenomena.
- Draft of short publication and thesis.

Are you working essentially on your own or as part of a team? If you are part of a team what is your role, and to what extent is the success of your project dependent on other members of the team?

I will be working as part of a team, and collaborate with other members for the project development, however my project will be focused on a particular direction and independent from the rest of the team progress.

What resources (equipment, materials, technician support etc.) will you need?

PC, Access of workstation and data server (supply by the group), maybe GPU cluster at ARC.

Do you require any training to meet your objectives, e.g. in the use of specific experimental equipment or software, and how are you going to obtain that training?

No. Coding skills in MATLAB and python are required, but these will be developed during the project.

Complete the following plan for your entire project as you see it now. List each major task down the left hand column, and for each one draw a horizontal line to indicate the period you expect to allocate to it. For example, the final task, writing your thesis, is shown as occupying mid-April to mid-June.

Task	Oct	Nov	Dec	Jan	Feb	Mar	Apr	May	Jun
Set up account on PSG group workstation server	xxxx								

Accessing of training data, learning how to use existing MATLAB code to prepare training data.	xxxx	xxxx						
Learning how to use PyTorch and produce effective deep learning models.	xxxx	xxxx						
Development, training, and testing of deep learning model		xxxx	xxxx	xxxx				
Selection of data set to be analysed for a particular scientific application.				xxxx	xxxx			
Analysis of data and quantification of scientific phenomena					xxxx	xxxx		
Drafting short publication (if possible)								
Writing up thesis							xxxx	xxxx

Has your supervisor completed a **Risk Assessment Form** about your project yet? Yes No

Your signature: *R. A. Horaan*
Date: 12/09/23

Your supervisor's signature: *P. M. J. Hafftka*
Date: *12/01/2023*

PROJECT MANAGEMENT FORM 2

1st Part II Project Analysis Form

Complete this form and send a copy to the Undergraduate Studies Office via
undergraduate.studies@materials.ox.ac.uk by Friday of 6th week of Michaelmas Term

Name: Rohit Abraham

Title of Project as given in your Project Description: Artificial Intelligence for multi-modal X-ray Imaging of metal solidification.

Refer back to the project plan in your Project Description and list the goals you set for this term. Comment briefly on the extent to which you have achieved them.

- For this term, the main goals were:
 - o Setting up access to PSG file servers and workstations
 - o Accessing and extracting the initial training data and labels
 - o Learning how to use PyTorch and producing an initial deep learning model.
 - o Using the prepared data to start training a specific object detection model for the aluminium grain data.
- Access to technology and files has been completed fine. Worked closely with Enzo to use certain packages and files which have not been used for 2-3 years. Plan to eventually help set up clear tutorials / brief learning materials for internal group usage.
- Extracting training data has been occurring over the last few weeks. Due to the nature of the task, it has been a time intensive process.
- I have received guidance on gathering the labels for the data. Currently working with both Enzo and Carlos Arteta on methods to effectively move the labels from their current format in MATLAB to one which can be used in modern Python based deep learning models.
- Completed initial tutorials in PyTorch and elementary deep learning techniques that have been provided by Andrew Zisserman's group. Will be developing these skills further when data preparation has been completed.

Identify clearly any difficulties you have encountered. Are they surmountable in the time available?

- Extraction of data from current file handling systems (MATLAB based) has taken quite a large amount of time. This has been due to the large nature of the dataset (>50GB) and low-level graphics issues within MATLAB leading to close supervision of the running code. However, the project progress should ensure that the data extraction and label retrieval is the final task in the project which should primarily rely on MATLAB, after which the majority of the project will occur on a Python based environment.
- Python environments within the workstation have produced some issues due to software incompatibility / hardware inefficiencies in the current workstation setup. The current workstation is required to be used by multiple users across the group and hence has been set up to be optimal for all users, which leads to some issues for specific machine learning tasks. This has been mitigated by generous investment from Enzo Liotti into a new linux-based workstation specifically for the development of deep learning models within the group to be used by myself and DPhil student Shun Yang. This will enable faster development of large object detection models. This workstation is being delivered and set up over the next few weeks.

State any refinements, modifications or replacements of the original objectives for your Part II project:

The objectives remain the same, however we have refined the exact scope of the project and identified further work if additional time is available later this year. It is outlined below:

- Setting up and using PyTorch to develop and train a CNN (Mask R-CNN) model to analyse the high quality XRT experiments.
- Count the number of equiaxed grains to a similar level or better than the existing MATLAB models. This will include experiments both with electromagnetic stirring and without.
- Track the movement of grains, how do they move and what are their paths of movement?
- Track the growth of the grains through the mask feature of the model.
- Understand growth effects through relating growth with local composition. Specifically, through the context of grain nucleation in waves, as demonstrated in *Enzo Liotti et al. Crystal nucleation in metallic alloys using x-ray radiography and machine learning (2018)*.

Further work includes:

- Applying and verifying the model with lower quality lab gathered XRT data.
- Looking further into the solidification processes and points where the eutectic transformation takes place.

Are you intending to change the title of your project? If so, state the new title:

No

Have the training needs you identified in the Project Description been met, and have you identified any further training requirements?

Training has been provided at any points where necessary.

Tick the appropriate box. Do you have:

Results	<input checked="" type="checkbox"/> None	<input type="checkbox"/> Some	<input type="checkbox"/> Sufficient
Analysis of results	<input checked="" type="checkbox"/> None	<input type="checkbox"/> Some	<input type="checkbox"/> Sufficient

Do you have any other comments you wish to make?

No

After looking at the project plan in your Project Description complete the following project plan for the remainder of your Part II.

Task	Dec	Jan	Feb	Mar	Apr	May	Jun
Transfer of training data and labels over to new workstation	xxxx						
Initial development of first deep learning model	xxxx	xxxx					
Refinement of models and feedback from Andrew / Carlos		xxxx					
Selection of new data and extraction of quantitative data on grains		xxxx	xxxx				
Development of further work / publishing work to open-source repositories (GitHub etc.)				xxxx	xxxx		
Drafting short publication							
Writing up					xxxxx	xxxxx	xxxxx

General comments by the supervisor:

R. Abraham

Your signature:
Date: 19/11/2023

Enzo Liotti

Your supervisor's signature:
Date: 20/11/2023

PROJECT MANAGEMENT FORM 3

2nd Part II Project Analysis Form

Complete this form and send a copy to the Undergraduate Studies Office via undergraduate.studies@materials.ox.ac.uk by Friday of 6th week of Hilary Term

Name: Rohit Abraham

Title of Project: Artificial Intelligence for X-ray Imaging of metal solidification.

Refer back to the project plan you made last term and list the goals you set for this term. Comment briefly on the extent to which you have achieved them.

- At the midpoint of Hilary term, I have been able to properly extract all previously annotated data from MATLAB based file handling systems and created records to filter higher quality and remove lower quality data.
- First few attempts at training Faster R-CNN based deep learning models have shown success but further work is needed to improve model performance, specifically with refining datasets and optimising hyperparameters.
- The movement and growth of crystals have been visually identified by the model, but further work is required to quantitatively gather these data points.
- Producing masks and segmentation has not been worked on, but instead identified as a task for further work once a completed and functional model has been produced to conduct the first stage of analysis on grain counting and motion.

Identify clearly any difficulties you have encountered. Are they surmountable in the time available?

- Around half of my data was incorrectly exported, leading to loss of synchronisation of annotations. This led to a set-back in model development as extra time was spent fixing these issues. Help from Enzo and Carlos has enabled this to be fixed and now we have a much larger database to train the model with.
- Delays in the procurement of the linux-based workstation has led to software / technological uncertainties. Certain software packages and tools that are unavailable on the Windows based machine has led to additional time spent writing scripts to manually produce equivalent tools.

State any refinements, modifications or replacements of the objectives you set for your Part II project:

- The primary focus currently is to produce a model which can accurately count the number of grains at a given time in a solidification experiment. Once this is complete, I will be looking to employ this model to conduct analysis over selected experiments to produce the bulk scientific content for my thesis.
- Once this is complete, under the guidance of my supervisors, we will be looking towards how we can use other deep learning architectures to further expand the applications of a crystal detection model, while also refining the work we have already completed.

Are you intending to change the title of your project?

Yes. If so, state the new title:

No

What is the title of the talk you will give to the Department?

Understanding the effect of convection on nucleation of equiaxed grains in Al-Cu alloys using Deep Learning and X-ray radiography

Have all your training needs for this project now been met?

Training and guidance has been provided by both Enzo Liotti and Carlos Arteta where necessary, any other training needs are self-taught.

Tick the appropriate box. Do you have:

Results	<input type="checkbox"/> None	<input checked="" type="checkbox"/> Some	<input type="checkbox"/> Sufficient
Analysis of results	<input checked="" type="checkbox"/> None	<input type="checkbox"/> Some	<input type="checkbox"/> Sufficient

Do you have any other comments you wish to make?

No

General comments by the supervisor:

The project is progressing well, although with some setbacks, and will quickly produce some interesting results once the training of the deep learning models is finished.

Your signature: *R. Alvaro*
Date: 16/04/24

Your supervisor's signature: *Enzo Liotti*
Date: 16/02/24

7.3 Health and Safety

Since this project was computational, there were limited risks but still remained an important consideration. The main identified risks derived from using a computer for extended periods of time. The majority of my work was conducted from home as all the systems being used would be used via. remote connection regardless of my location.

Due to previous work placements being conducted in remote capacities, I had received mandatory training on ensuring an ergonomic desk set up to avoid injury. A monitor was used with a laptop and adjusted at an appropriate height to avoid neck strain.

I made it a priority to remain active and take breaks during work, especially while waiting for longer tasks to be complete during the day. When working on tasks outside of model development and analysis, such as background reading and thesis writing, I aimed to spend portions of the day in a library or public space for general wellness.

During the early stage of the project, it was anticipated I may have some contribution towards future radiography experiments. I attended training on the Safe Use of X-ray Generators, Sealed Sources & Accelerators provided by the University of Oxford Safety Office. While this was not needed in the end, it was an informative presentation and provided useful background.

Appendix A

Appendix

A.1 Evolution of Region-Based CNN Architectures

Faster R-CNN, employed in this project to develop the grain detection model, builds upon prior work by Girshick *et al.* (2013) who proposed the R-CNN network [31] and followed with the Fast R-CNN in 2015 [74]. A brief outline of how the two models improved upon each other leading to Faster R-CNN has been provided.

A.1.1 R-CNN

R-CNN is designed to propose regions of interest where an object may exist and then use a CNN to classify each of the regions. The method in which the regions are proposed strongly influences the accuracy of the network [66, 101], and for R-CNN this was initially done via. a selective search approach [73]. For classification, CNNs such as AlexNet, first introduced by Krizhevsky *et al.* in 2012 [28] or VGG-16, first introduced by Simonyan & Zisserman in 2014 [29] were employed. Since the CNN is used to extract features from each proposed region independently, the overall network was computationally expensive to run. The CNN used is typically referred to as backbone of the model.

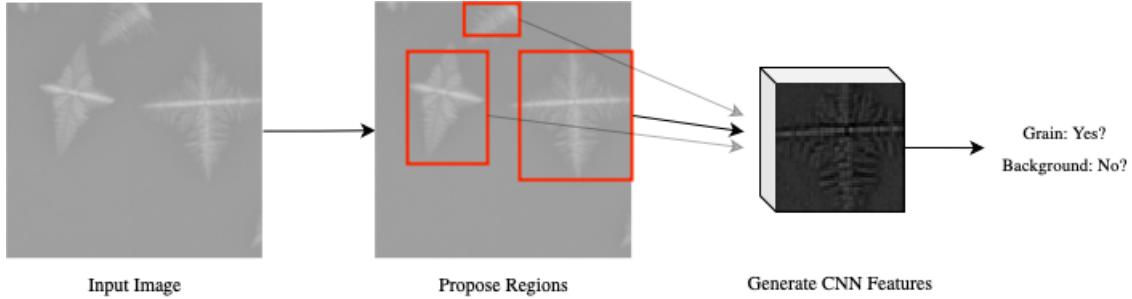


Figure A.1: R-CNN network overview (adapted from Girshick *et al.* (2014) [31])

A.1.2 Fast R-CNN

Detailed by Girshick in 2015, Fast R-CNN aimed to improve on the computational inefficiencies of R-CNN by reducing the number of times a CNN was employed to extract features from an image. Selective search methods are still used to propose regions, however instead of processing each region through the CNN, the CNN is used once to produce a convolutional feature map for the entire image. The feature map and the object proposals are combined to create regions of interest, formally detailed as a Region of Interest (RoI) pooling layer and respective feature vectors. The model then determines probabilities that a specific region belongs to each class, which would be either grain or background for this application, alongside the coordinates of the bounding box [74]. Using this method significantly increased computational efficiency, however selective search is still relied on to propose regions and represents a bottleneck in both speed and accuracy of proposals.

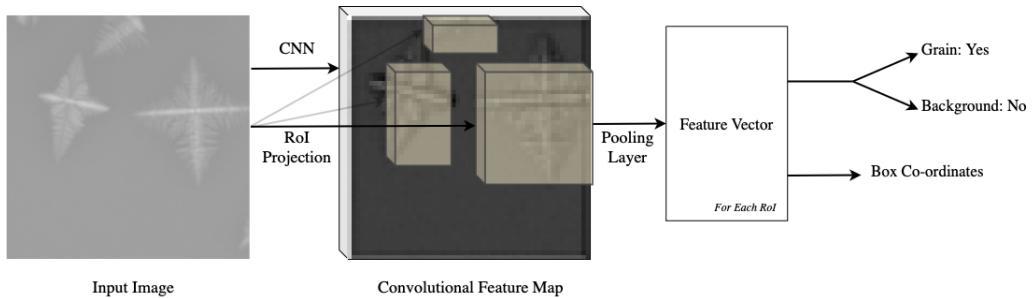


Figure A.2: Fast R-CNN network overview (adapted from Girshick (2015) [74])

References

- [1] IEA, *Iron and Steel Technology Roadmap*, <https://www.iea.org/reports/iron-and-steel-technology-roadmap>, IEA, Paris. Licence: CC BY 4.0, 2020.
- [2] A. Graham, M. Ford, J. Martins, and M. P. Partnership, “Aluminum decarbonization at a cost that makes sense,” 2022, Analytical support from McKinsey & Company. [Online]. Available: <https://missionpossiblepartnership.com>.
- [3] I. R. P. United Nations Environment Programme, *Recycling rates of metals: A status report*, 2011. [Online]. Available: <https://wedocs.unep.org/20.500.11822/8702>.
- [4] World Steel Association, “Climate change and the production of iron and steel,” 2021. [Online]. Available: <https://worldsteel.org/publications/policy-papers/climate-change-policy-paper/>.
- [5] Y. He, K. Zhou, Y. Zhang, H. Xiong, and L. Zhang, “Recent progress of inert anodes for carbon-free aluminium electrolysis: A review and outlook,” *Journal of Materials Chemistry A*, vol. 9, pp. 25 272–25 285, 45 2021. DOI: 10.1039/d1ta07198j.
- [6] M. Rappaz and J. A. Dantzig, *Solidification*. EPFL Press, Aug. 2009, ISBN: 0849382386.
- [7] C. Kenel and C. Leinenbach, “Influence of cooling rate on microstructure formation during rapid solidification of binary tial alloys,” *Journal of Alloys and Compounds*, vol. 637, pp. 242–247, 2015. DOI: 10.1016/j.jallcom.2015.03.016.

REFERENCES

- [8] J. Jezierski, M. Dojka, M. Stawarz, and R. Dojka, “Scanning electron microscopy as a tool for castings quality analysis,” *Archives of Foundry Engineering*, vol. 22, no. 1, pp. 53–61, 2022.
- [9] D. Zhang *et al.*, “Microstructure evolution and enhanced mechanical properties in al-mn alloy reinforced by b-doped tic particles,” *Materials & Design*, vol. 221, p. 110906, 2022. DOI: [10.1016/j.matdes.2022.110906](https://doi.org/10.1016/j.matdes.2022.110906).
- [10] X. Di and B. Chen, “Solidification behaviour of a high chromium nickel base alloy weld deposited metal,” *Science and Technology of Welding and Joining*, vol. 20, pp. 325–329, 4 2015. DOI: [10.1179/1362171815y.0000000019](https://doi.org/10.1179/1362171815y.0000000019).
- [11] S. Karagadde, C. L. A. Leung, and P. Lee, “Progress on in situ and operando x-ray imaging of solidification processes,” *Materials*, vol. 14, p. 2374, 9 2021. DOI: [10.3390/ma14092374](https://doi.org/10.3390/ma14092374).
- [12] Y. Wang, S. Jia, M. Wei, L. Peng, Y. Wu, and X. Liu, “Research progress on solidification structure of alloys by synchrotron x-ray radiography: A review,” *Journal of Magnesium and Alloys*, vol. 8, pp. 396–413, 2 2020. DOI: [10.1016/j.jma.2019.08.003](https://doi.org/10.1016/j.jma.2019.08.003).
- [13] E. Liotti, C. Arteta, A. Zisserman, A. Lui, V. Lempitsky, and P. S. Grant, “Crystal nucleation in metallic alloys using x-ray radiography and machine learning,” *Science Advances*, vol. 4, 4 2018. DOI: [10.1126/sciadv.aar4004](https://doi.org/10.1126/sciadv.aar4004).
- [14] J. Epp, “4 - x-ray diffraction (xrd) techniques for materials characterization,” in *Materials Characterization Using Nondestructive Evaluation (NDE) Methods*, G. Hübschen, I. Altpeter, R. Tschuncky, and H.-G. Herrmann, Eds., Woodhead Publishing, 2016, pp. 81–124, ISBN: 978-0-08-100040-3. DOI: <https://doi.org/10.1016/B978-0-08-100040-3.00004-3>. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/B9780081000403000043>.

REFERENCES

- [15] S. Ebnesajjad, “Chapter 4 - surface and material characterization techniques,” in *Surface Treatment of Materials for Adhesive Bonding (Second Edition)*, S. Ebnesajjad, Ed., Second Edition, Oxford: William Andrew Publishing, 2014, pp. 39–75, ISBN: 978-0-323-26435-8. DOI: <https://doi.org/10.1016/B978-0-323-26435-8.00004-6>. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/B9780323264358000046>.
- [16] Y. Peng *et al.*, “Recent progress of synchrotron x-ray imaging and diffraction on the solidification and deformation behavior of metallic materials,” *Acta Metallurgica Sinica (English Letters)*, vol. 35, pp. 3–24, 1 2021. DOI: [10.1007/s40195-021-01311-4](https://doi.org/10.1007/s40195-021-01311-4).
- [17] Y. Wang *et al.*, “Revealing the diversity of dendritic morphology evolution during solidification of magnesium alloys using synchrotron x-ray imaging: A review,” *Acta Metallurgica Sinica (English Letters)*, vol. 35, pp. 177–200, 2 2021. DOI: [10.1007/s40195-021-01319-w](https://doi.org/10.1007/s40195-021-01319-w).
- [18] S. Feng, E. Liotti, and P. S. Grant, “X-ray imaging of alloy solidification: Crystal formation, growth, instability and defects,” *Materials*, vol. 15, p. 1319, 4 2022. DOI: [10.3390/ma15041319](https://doi.org/10.3390/ma15041319).
- [19] H. Bale *et al.*, “Real-time quantitative imaging of failure events in materials under load at temperatures above 1,600 °c,” *Nature Materials*, vol. 12, pp. 40–46, 1 2012. DOI: [10.1038/nmat3497](https://doi.org/10.1038/nmat3497).
- [20] S. Feng, E. Liotti, M. D. Wilson, L. Jowitt, and P. S. Grant, “In situ mapping of chemical segregation using synchrotron x-ray imaging,” *MRS Bulletin*, vol. 45, pp. 934–942, 11 2020. DOI: [10.1557/mrs.2020.270](https://doi.org/10.1557/mrs.2020.270).
- [21] N. Wang *et al.*, “Dynamic evolution of microstructure morphology in thin-sample solidification: Deep learning assisted synchrotron x-ray radiography,” *Materials*

REFERENCES

- Characterization*, vol. 181, p. 111451, 2021. DOI: 10.1016/j.matchar.2021.111451.
- [22] C. M. Bishop, *Pattern Recognition and Machine Learning*, 1st ed. Springer, 2006, ISBN: 1493938436.
- [23] I. H. Sarker, “Deep learning: A comprehensive overview on techniques, taxonomy, applications and research directions,” *SN Computer Science*, vol. 2, 6 2021. DOI: 10.1007/s42979-021-00815-1.
- [24] T. Ridnik, H. Lawen, A. Noy, E. B. Baruch, G. Sharir, and I. Friedman, “Tresnet: High performance gpu-dedicated architecture,” 2020. DOI: 10.48550/arxiv.2003.13630.
- [25] D. Amodei, D. Hernandez, G. Sastry, J. Clark, G. Brockman, and I. Sutskever, *Ai and compute*, May 2018. [Online]. Available: <https://openai.com/index/ai-and-compute/>.
- [26] P. A. Viola and M. Jones, “Rapid object detection using a boosted cascade of simple features,” *Proceedings of the 2001 IEEE Computer Society Conference on Computer Vision and Pattern Recognition. CVPR 2001*, DOI: 10.1109/cvpr.2001.990517.
- [27] 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, 2005, 886–893 vol. 1. DOI: 10.1109/CVPR.2005.177.
- [28] A. Krizhevsky, I. Sutskever, and G. E. Hinton, “Imagenet classification with deep convolutional neural networks,” *Communications of the ACM*, vol. 60, pp. 84–90, 6 2017. DOI: 10.1145/3065386.
- [29] K. Simonyan and A. Zisserman, *Very deep convolutional networks for large-scale image recognition*, 2015. arXiv: 1409.1556 [cs.CV].
- [30] K. He, X. Zhang, S. Ren, and J. Sun, *Deep residual learning for image recognition*, 2015. arXiv: 1512.03385 [cs.CV].

REFERENCES

- [31] R. Girshick, J. Donahue, T. Darrell, and J. Malik, *Rich feature hierarchies for accurate object detection and semantic segmentation*, 2014. arXiv: 1311.2524 [cs.CV].
- [32] P. Mikołajczak, “Microstructural evolution in almgisi alloys during solidification under electromagnetic stirring,” *Metals*, vol. 7, p. 89, 3 2017. DOI: 10.3390/met7030089.
- [33] J. Spittle, “Grain refinement in shape casting of aluminium alloys,” *International Journal of Cast Metals Research*, vol. 19, pp. 210–222, 4 2006. DOI: 10.1179/136404606225023444.
- [34] A. Hamasaaid *et al.*, “Effect of mold coating materials and thickness on heat transfer in permanent mold casting of aluminum alloys,” *Metallurgical and Materials Transactions A*, vol. 38, pp. 1303–1316, 6 2007. DOI: 10.1007/s11661-007-9145-2.
- [35] S. Li *et al.*, “Development and applications of aluminum alloys for aerospace industry,” *Journal of Materials Research and Technology*, vol. 27, pp. 944–983, 2023. DOI: 10.1016/j.jmrt.2023.09.274.
- [36] T. A. Association, *International alloy designations and chemical composition limits for wrought aluminum and wrought aluminum alloys*, Jan. 2015.
- [37] B. Bellón *et al.*, “Multiscale prediction of microstructure length scales in metallic alloy casting,” *Acta Materialia*, vol. 207, p. 116686, 2021. DOI: 10.1016/j.actamat.2021.116686.
- [38] S. Feng *et al.*, “An in-situ method to estimate the tip temperature and phase selection of secondary fe-rich intermetallics using synchrotron x-ray radiography,” *Scripta Materialia*, vol. 149, pp. 44–48, 2018. DOI: 10.1016/j.scriptamat.2018.02.001.

REFERENCES

- [39] E. O. Huerta, A. Oliva, F. Avilés, J. González-Hernández, and J. Corona, “Elastic modulus determination of al-cu film alloys prepared by thermal diffusion,” *Journal of Nanomaterials*, vol. 2012, pp. 1–8, 2012. DOI: [10.1155/2012/895131](https://doi.org/10.1155/2012/895131).
- [40] B. Murty, S. Kori, and M. Chakraborty, “Grain refinement of aluminium and its alloys by heterogeneous nucleation and alloying,” *International Materials Reviews*, vol. 47, pp. 3–29, 1 2002. DOI: [10.1179/095066001225001049](https://doi.org/10.1179/095066001225001049).
- [41] S. Henry, T. Minghetti, and M. Rappaz, “Dendrite growth morphologies in aluminium alloys,” *Acta Materialia*, vol. 46, pp. 6431–6443, 18 1998. DOI: [10.1016/s1359-6454\(98\)00308-5](https://doi.org/10.1016/s1359-6454(98)00308-5).
- [42] S. Feng, E. Liotti, A. Lui, M. D. Wilson, and P. S. Grant, “Nucleation bursts of primary intermetallic crystals in a liquid al alloy studied using in situ synchrotron x-ray radiography,” *Acta Materialia*, vol. 221, p. 117389, 2021. DOI: [10.1016/j.actamat.2021.117389](https://doi.org/10.1016/j.actamat.2021.117389).
- [43] A. Prasad, S. D. McDonald, H. Yasuda, K. Nogita, and D. H. StJohn, “A real-time synchrotron x-ray study of primary phase nucleation and formation in hypoeutectic al-si alloys,” *Journal of Crystal Growth*, vol. 430, pp. 122–137, 2015. DOI: [10.1016/j.jcrysgro.2015.06.024](https://doi.org/10.1016/j.jcrysgro.2015.06.024).
- [44] A. Prasad *et al.*, “Real-time synchrotron x-ray observations of equiaxed solidification of aluminium alloys and implications for modelling,” *IOP Conference Series: Materials Science and Engineering*, vol. 84, p. 012014, 2015. DOI: [10.1088/1757-899x/84/1/012014](https://doi.org/10.1088/1757-899x/84/1/012014).
- [45] D. H. StJohn, M. Qian, M. Easton, and P. Cao, “The interdependence theory: The relationship between grain formation and nucleant selection,” *Acta Materialia*, vol. 59, pp. 4907–4921, 12 2011. DOI: [10.1016/j.actamat.2011.04.035](https://doi.org/10.1016/j.actamat.2011.04.035).
- [46] M. Safari, M. Keikha, and A. Kamarei, “An investigation on semi-solid forming of a360 aluminium alloy by mechanical stirring,” *Proceedings of the Institution of*

REFERENCES

- Mechanical Engineers, Part E: Journal of Process Mechanical Engineering*, vol. 226, pp. 205–215, 3 2011. DOI: [10.1177/0954408911411808](https://doi.org/10.1177/0954408911411808).
- [47] N. Barman and P. Dutta, “Effect of process parameters on transport phenomena during solidification in the presence of electromagnetic stirring,” *Transactions of the Indian Institute of Metals*, vol. 62, pp. 469–474, 4-5 2009. DOI: [10.1007/s12666-009-0063-4](https://doi.org/10.1007/s12666-009-0063-4).
- [48] B. Ren, D. Chen, H. Wang, and M. Long, “Numerical analysis of coupled turbulent flow and macroscopic solidification in a round bloom continuous casting mold with electromagnetic stirring,” *Steel Research International*, vol. 86, pp. 1104–1115, 9 2015. DOI: [10.1002/srin.201400178](https://doi.org/10.1002/srin.201400178).
- [49] J. Li, B. Wang, Y. Ma, and J. Chen, “Effect of complex electromagnetic stirring on inner quality of high carbon steel bloom,” *Materials Science and Engineering: A*, vol. 425, pp. 201–204, 1-2 2006. DOI: [10.1016/j.msea.2006.03.061](https://doi.org/10.1016/j.msea.2006.03.061).
- [50] L. B. Trindade, J. E. A. Nadalon, A. C. Contini, and R. C. Barroso, “Modeling of solidification in continuous casting round billet with mold electromagnetic stirring (m-ems),” *Steel Research International*, vol. 88, p. 1600319, 4 2016. DOI: [10.1002/srin.201600319](https://doi.org/10.1002/srin.201600319).
- [51] D. Jiang and M. Zhu, “Flow and solidification in billet continuous casting machine with dual electromagnetic stirrings of mold and the final solidification,” *Steel Research International*, vol. 86, pp. 993–1003, 9 2014. DOI: [10.1002/srin.201400281](https://doi.org/10.1002/srin.201400281).
- [52] Z. Zhang, X. Chen, J. Xu, and S. Li-kai, “Numerical simulation on electromagnetic field, flow field and temperature field in semisolid slurry preparation by a-ems,” *Rare Metals*, vol. 29, pp. 635–641, 6 2010. DOI: [10.1007/s12598-010-0184-2](https://doi.org/10.1007/s12598-010-0184-2).
- [53] D. Jiang and M. Zhu, “Flow and solidification in billet continuous casting machine with dual electromagnetic stirrings of mold and the final solidification,” *Steel Research International*, vol. 86, pp. 993–1003, 9 2014. DOI: [10.1002/srin.201400281](https://doi.org/10.1002/srin.201400281).

REFERENCES

- [54] B. Willers *et al.*, “Efficient melt stirring using pulse sequences of a rotating magnetic field: Part ii. application to solidification of al-si alloys,” *Metallurgical and Materials Transactions B*, vol. 39, pp. 304–316, 2 2008. DOI: [10.1007/s11663-008-9126-x](https://doi.org/10.1007/s11663-008-9126-x).
- [55] O. Bustos, R. Leiva, and C. Sánchez, “Effect of magnetic stirring on alfeSi intermetalics of an a356 aluminum alloy,” *Matéria (Rio De Janeiro)*, vol. 27, 2 2022. DOI: [10.1590/1517-7076-rmat-2022-48310](https://doi.org/10.1590/1517-7076-rmat-2022-48310).
- [56] E. Liotti *et al.*, “A synchrotron x-ray radiography study of dendrite fragmentation induced by a pulsed electromagnetic field in an al-15cu alloy,” *Acta Materialia*, vol. 70, pp. 228–239, 2014, ISSN: 1359-6454. DOI: <https://doi.org/10.1016/j.actamat.2014.02.024>. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S1359645414001177>.
- [57] E. Liotti *et al.*, “The spatial and temporal distribution of dendrite fragmentation in solidifying al-cu alloys under different conditions,” *Acta Materialia*, vol. 121, pp. 384–395, 2016, ISSN: 1359-6454. DOI: <https://doi.org/10.1016/j.actamat.2016.09.013>. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S1359645416307017>.
- [58] S. Bhagavath *et al.*, “Combined deformation and solidification-driven porosity formation in aluminum alloys,” *Metallurgical and Materials Transactions A*, vol. 50, pp. 4891–4899, 10 2019. DOI: [10.1007/s11661-019-05378-8](https://doi.org/10.1007/s11661-019-05378-8).
- [59] I. Arganda-Carreras *et al.*, “Trainable weka segmentation: A machine learning tool for microscopy pixel classification,” *Bioinformatics*, vol. 33, pp. 2424–2426, 15 2017. DOI: [10.1093/bioinformatics/btx180](https://doi.org/10.1093/bioinformatics/btx180).
- [60] J. Schindelin *et al.*, “Fiji: An open-source platform for biological-image analysis,” *Nature Methods*, vol. 9, pp. 676–682, 7 2012. DOI: [10.1038/nmeth.2019](https://doi.org/10.1038/nmeth.2019).

REFERENCES

- [61] C. S. Kaira *et al.*, “Automated correlative segmentation of large transmission x-ray microscopy (txm) tomograms using deep learning,” *Materials Characterization*, vol. 142, pp. 203–210, 2018. DOI: 10.1016/j.matchar.2018.05.053.
- [62] Z. Li *et al.*, “Characterization of the convoluted 3d intermetallic phases in a recycled al alloy by synchrotron x-ray tomography and machine learning,” *Acta Metallurgica Sinica (English Letters)*, vol. 35, pp. 115–123, 1 2021. DOI: 10.1007/s40195-021-01312-3.
- [63] J. Yeom, T. Stan, S. Hong, and P. W. Voorhees, “Segmentation of experimental datasets via convolutional neural networks trained on phase field simulations,” *Acta Materialia*, vol. 214, p. 116990, 2021. DOI: 10.1016/j.actamat.2021.116990.
- [64] C. T. C. Inc., *Workflow and data preparation*, 2020. [Online]. Available: <http://www.theobjects.com/dragonfly/dfhelp/2020-1/Content/Artificial%20Intelligence/Deep%20Learning%20Tool/Workflow%20and%20Data%20Preparation.htm>.
- [65] A. R. Rafael Padilla and the Hugging Face Team, *Open object detection leaderboard*, https://huggingface.co/spaces/rafaelpadilla/object_detection_leaderboard, 2023.
- [66] S. Ren, K. He, R. Girshick, and J. Sun, *Faster r-cnn: Towards real-time object detection with region proposal networks*, 2016. arXiv: 1506.01497 [cs.CV].
- [67] Y. Zhu, F. Wang, Q. Chen, and L. Li, “Research on the application of object detection in the installation of spacer bars,” *Journal of Physics Conference Series*, vol. 2303, p. 012075, 1 2022. DOI: 10.1088/1742-6596/2303/1/012075.
- [68] Q. Guo, L. Liu, W. Xu, Y. Gong, X. Zhang, and W. Jing, “An improved faster r-cnn for high-speed railway dropper detection,” *Ieee Access*, vol. 8, pp. 105 622–105 633, 2020. DOI: 10.1109/access.2020.3000506.

REFERENCES

- [69] M. Musci, C. Persello, and A. Lingua, “Uav images and deep-learning algorithms for detecting flavescence doree disease in grapevine orchards,” *The International Archives of the Photogrammetry Remote Sensing and Spatial Information Sciences*, vol. XLIII-B3-2020, pp. 1483–1489, 2020. DOI: 10.5194/isprs-archives-xliii-b3-2020-1483-2020.
- [70] J. Redmon, S. Divvala, R. Girshick, and A. Farhadi, *You only look once: Unified, real-time object detection*, 2016. arXiv: 1506.02640 [cs.CV].
- [71] W. Liu *et al.*, “Ssd: Single shot multibox detector,” in *Lecture Notes in Computer Science*. Springer International Publishing, 2016, pp. 21–37, ISBN: 9783319464480. DOI: 10.1007/978-3-319-46448-0_2. [Online]. Available: http://dx.doi.org/10.1007/978-3-319-46448-0_2.
- [72] A. Paszke *et al.*, *Pytorch: An imperative style, high-performance deep learning library*, 2019. arXiv: 1912.01703 [cs.LG].
- [73] J. Uijlings, K. E. A. v. d. Sande, T. Gevers, and A. Smeulders, “Selective search for object recognition,” *International Journal of Computer Vision*, vol. 104, pp. 154–171, 2 2013. DOI: 10.1007/s11263-013-0620-5.
- [74] R. Girshick, *Fast r-cnn*, 2015. arXiv: 1504.08083 [cs.CV].
- [75] S. Raschka and V. Mirjalili, *Python machine learning: Machine learning and deep learning with Python, scikit-learn, and TensorFlow 2*. Packt publishing ltd, 2019.
- [76] J. Ansel *et al.*, “Pytorch 2: Faster machine learning through dynamic python bytecode transformation and graph compilation,” in *Proceedings of the 29th ACM International Conference on Architectural Support for Programming Languages and Operating Systems, Volume 2*, ser. ASPLOS ’24, ¶conf-loc¶, ¶city¶La Jolla¶/city¶, ¶state¶CA¶/state¶, ¶country¶USA¶/country¶, ¶/conf-loc¶: Association for Computing Machinery, 2024, pp. 929–947, ISBN: 9798400703850. DOI: 10.1145/3620665.3640366. [Online]. Available: <https://doi.org/10.1145/3620665.3640366>.

REFERENCES

- [77] N. Corporation, *What is pytorch?* 2024. [Online]. Available: <https://www.nvidia.com/en-us/glossary/pytorch/>.
- [78] T. maintainers and contributors, *Torchvision: Pytorch's computer vision library*, <https://github.com/pytorch/vision>, 2016.
- [79] A. Dutta and A. Zisserman, “The VIA annotation software for images, audio and video,” in *Proceedings of the 27th ACM International Conference on Multimedia*, ser. MM ’19, Nice, France: ACM, 2019, ISBN: 978-1-4503-6889-6/19/10. DOI: 10.1145/3343031.3350535. [Online]. Available: <https://doi.org/10.1145/3343031.3350535>.
- [80] A. Dutta and D. Miguel Susano Pinto, *List annotator (lisa)*, <https://gitlab.com/vgg/lisa>, 2020.
- [81] A. Buslaev, V. I. Iglovikov, E. Khvedchenya, A. Parinov, M. Druzhinin, and A. A. Kalinin, *A list of transforms and their supported targets*, 2020. [Online]. Available: https://albumentations.ai/docs/getting_started/transforms_and_targets/.
- [82] B. Zoph, E. Cubuk, G. Ghiasi, T. Lin, J. Shlens, and Q. Le, “Learning data augmentation strategies for object detection,” 2019. DOI: 10.48550/arxiv.1906.11172.
- [83] N. Tran, V. Tran, N. Nguyen, T. Nguyen, and N. Cheung, “On data augmentation for gan training,” *IEEE Transactions on Image Processing*, vol. 30, pp. 1882–1897, 2021. DOI: 10.1109/tip.2021.3049346.
- [84] T. DeVries and G. W. Taylor, *Dataset augmentation in feature space*, 2017. arXiv: 1702.05538 [stat.ML].
- [85] L. Perez and J. Wang, *The effectiveness of data augmentation in image classification using deep learning*, 2017. arXiv: 1712.04621 [cs.CV].
- [86] A. Buslaev, V. I. Iglovikov, E. Khvedchenya, A. Parinov, M. Druzhinin, and A. A. Kalinin, “Albumentations: Fast and flexible image augmentations,” *Information*,

REFERENCES

- vol. 11, no. 2, 2020, ISSN: 2078-2489. DOI: 10.3390/info11020125. [Online]. Available: <https://www.mdpi.com/2078-2489/11/2/125>.
- [87] J. Deng, W. Dong, R. Socher, L.-J. Li, K. Li, and L. Fei-Fei, “Imagenet: A large-scale hierarchical image database,” in *2009 IEEE Conference on Computer Vision and Pattern Recognition*, 2009, pp. 248–255. DOI: 10.1109/CVPR.2009.5206848.
- [88] S. Bozinovski and A. Fulgosi, “The influence of pattern similarity and transfer learning upon the training of a base perceptron b2,” Croatian, in *Proceedings of Symposium Informatica*, Original in Croatian, Bled, 1976, pp. 3-121–3-125.
- [89] T.-Y. Lin *et al.*, *Microsoft coco: Common objects in context*, 2015. arXiv: 1405.0312 [cs.CV].
- [90] Martín Abadi *et al.*, *TensorFlow: Large-scale machine learning on heterogeneous systems*, Software available from tensorflow.org, 2015. [Online]. Available: <https://www.tensorflow.org/>.
- [91] R. Padilla, W. L. Passos, T. L. B. Dias, S. L. Netto, and E. A. B. da Silva, “A comparative analysis of object detection metrics with a companion open-source toolkit,” *Electronics*, vol. 10, no. 3, 2021, ISSN: 2079-9292. DOI: 10.3390/electronics10030279. [Online]. Available: <https://www.mdpi.com/2079-9292/10/3/279>.
- [92] R. Padilla, S. L. Netto, and E. A. B. da Silva, “A survey on performance metrics for object-detection algorithms,” in *2020 International Conference on Systems, Signals and Image Processing (IWSSIP)*, 2020, pp. 237–242.
- [93] M. Everingham, S. M. A. Eslami, L. V. Gool, C. K. I. Williams, J. Winn, and A. Zisserman, “The pascal visual object classes challenge: A retrospective,” *International Journal of Computer Vision*, vol. 111, pp. 98–136, 1 2014. DOI: 10.1007/s11263-014-0733-5.
- [94] N. Wojke, A. Bewley, and D. Paulus, *Simple online and realtime tracking with a deep association metric*, 2017. arXiv: 1703.07402 [cs.CV].

REFERENCES

- [95] K. He, G. Gkioxari, P. Dollár, and R. Girshick, *Mask r-cnn*, 2018. arXiv: 1703.06870 [cs.CV].
- [96] A. Kirillov *et al.*, *Segment anything*, 2023. arXiv: 2304.02643 [cs.CV].
- [97] D. Raabe *et al.*, “Making sustainable aluminum by recycling scrap: The science of “dirty” alloys,” *Progress in Materials Science*, vol. 128, p. 100947, 2022. DOI: 10.1016/j.pmatsci.2022.100947.
- [98] A. de Vries, “The growing energy footprint of artificial intelligence,” *Joule*, vol. 7, pp. 2191–2194, 10 2023. DOI: 10.1016/j.joule.2023.09.004.
- [99] S. Evans and V. Viisainen, *Analysis: Uk electricity from fossil fuels drops to lowest level since 1957*, Jan. 2024. [Online]. Available: <https://www.carbonbrief.org/analysis-uk-electricity-from-fossil-fuels-drops-to-lowest-level-since-1957/>.
- [100] Ofgem, *Average gas and electricity usage*, 2023. [Online]. Available: <https://www.ofgem.gov.uk/average-gas-and-electricity-usage>.
- [101] J. Hosang, R. Benenson, P. Dollar, and B. Schiele, “What makes for effective detection proposals?” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 38, no. 4, pp. 814–830, Apr. 2016, ISSN: 2160-9292. DOI: 10.1109/tpami.2015.2465908. [Online]. Available: <http://dx.doi.org/10.1109/TPAMI.2015.2465908>.