DEVELOPMENT OF ARTIFICIAL-INTELLIGENCE BASED AUTONOMOUS ROOF FALL HAZARD DETECTION SYSTEM

by

Ergin Isleyen

A thesis submitted to the Faculty and the Board of Trustees of the Colorado School of Mines in partial fulfillment of the requirements for the degree of Doctor of Philosophy (Mining and Earth Systems Engineering).

Golden, Colorado

Date ____________________

Signed: __________________________

Ergin Isleyen

Signed: __________________________

Dr. Sebnem H. Duzgun
Thesis Advisor

Signed: __________________________

Dr. McKell R. Carter
Thesis Co-Advisor

Golden, Colorado

Date ____________________

Signed: __________________________

Dr. M. Stephen Enders
Department Head
Department of Mining Engineering

ABSTRACT

Autonomous technologies have the potential to transform mining by improving efficiency, safety, and productivity. The transition to fully autonomous mining requires an interoperable system that is autonomous at all stages of a mining operation including geotechnical hazard management. Deep learning is a viable option to develop autonomous technologies based on the improvements in computing power and sensor technology. However, the success of deep learning development depends on having a large and diverse training set. Geotechnical hazard management requires on-site data collection. The restrictions on on-site data collection such as accessibility, time, and financial constraints cause small and unbalanced datasets that limit the performance of deep learning systems and thereby prevent the increased use of autonomous technologies. This thesis presents a methodology that includes transfer learning, synthetic data generation, and an image-feature based data sampling technique to train a convolutional neural network (CNN) using a small dataset collected on-site. The methodology is implemented in a case study to develop an autonomous roof fall hazard detection system. The case study is a large-opening limestone mine in the Midwestern United States that has frequent roof fall problems caused by high horizontal stresses. The typical hazard management approach for this type of roof fall hazard relies heavily on visual inspections and expert knowledge. Therefore, images depicting hazardous and non-hazardous roof conditions are collected based on expert labeling. Using the transfer learning approach, a pre-trained network provided the suitable low-level features to train the final fully connected layer of a CNN. The network classification accuracy was improved by expanding the training set with synthetic images rendered from a 3-D model generated with digital photogrammetry. Also, a data sampling technique that uses Gabor

magnitude responses was introduced. The objective of this technique is to improve the quality of the training set by removing non-informative samples, and it is shown that implementing this technique improves network performance. Besides, accuracy alone is not enough to ensure a reliable hazard management system. System constraints and reliability are improved when the features being used by the network are understood. Therefore, a deep learning interpretation technique called integrated gradients is used to identify the important features in each image for prediction. The analysis of integrated gradients shows that the system the same visual features as the expert on roof fall hazard detection. The final network performance is verified with a test dataset that had not been used during training. The original contribution of this research is a methodology to develop artificial intelligence-based, autonomous, and interpretable geotechnical hazard detection systems by implementing the techniques to overcome limited training data problem. The findings of this thesis demonstrate the potential of deep learning even with small and unbalanced datasets. The results provide a foundation for the increased use of autonomous technologies in situations in which the absence of large and diverse data previously prevented the development of deep learning systems.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF EQUATIONS

ACKNOWLEDGMENTS

I would like to express my wholehearted appreciation to my advisor, Dr. Sebnem Duzgun. I am forever grateful for the opportunity she provided to conduct this research. I am also grateful to my co-advisor, Dr. McKell Carter, for all his support and viewpoints that helped me refine my work. I would also like to extend my gratitude to my thesis committee members - Dr. Jurgen Brune, Dr. Jamal Rostami, and Dr. Thomas Williams - for their valuable suggestions.

Special thanks go to Tim Miller for providing access to Subtropolis Mine. His willingness to collaborate and his assistance during data collection improved the quality of this work significantly. Thank you to Robert Reeves for lending me the equipment used in data collection.

This research was partially supported by the National Institute for Occupational Safety and Health (NIOSH) contract number 0000HCCR-2019-36403.

And finally, I would like to thank the friends that I made during my time at Colorado School of Mines, especially Lukas Fahle, Gerardo Martinez, Hamilton Pinheiro, Natalia Soares Rodriguez, Dr. Hilal Soydan, Caleb Stetson and many others who made the last three years much more enjoyable.

CHAPTER 1

INTRODUCTION

Artificial intelligence (AI) and autonomous technologies have the potential to redefine many industries. While AI and autonomous technology applications are still in their early stages, utilizing the benefits of AI technologies will provide significant competitive advantages and immense productivity improvements (Makridakis, 2017). AI-based autonomous technologies have become a viable option for businesses due to an increase in the availability of high-performance computing power, and the improvements in sensors technology.

Autonomous technologies also have the potential to transform the mining industry by improving operational efficiency and safety performance, boosting productivity, and cutting costs. The mining industry has already started to see the impacts of autonomous technologies and robotics. Many mines throughout the world elevate their productivity and efficiency with autonomous haulage systems, and the number of companies with autonomous truck fleets is increasing. Drone technology makes it possible to survey large areas with high accuracy in record times without sending personnel into potentially dangerous situations. Robots replace humans to inspect small and inhospitable environments in mines. The World Economic Forum (2017) reported that, through autonomous operations and robotics, $47 billion of additional value can be created for the mining industry between 2016 – 2025. They also estimated that 120 lives will be saved and approximately 7000 injuries will be avoided within the same period worldwide. CO_2 emissions can be reduced by almost 340 million tonnes by using autonomous machines.

However, building fully autonomous mines remains a major challenge, despite its tremendous value. A report by Clareo and the World Innovation Network (2019) argues that closed and proprietary systems are the major barrier to implementation of fully autonomous mines, and the solution requires a shift to open and interoperable standards at all stages of a mining operation. Also, autonomous operations require mining companies to change how they manage workplace hazards. For example, the Department of Mines and Petroleum of Western Australia (2015) has already implemented a new code of practice that aims to promote improved hazard control for mines that use autonomous haulage systems. Based on these, it can be expected that the transition to fully autonomous mining will require a change in hazard management at a much larger scale.

Geotechnical hazard detection is a crucial part of mining operations. It includes assessing the likelihood, evaluating the impact, calculating the risk of the hazards such as roof falls, slope failures, and subsidence. In the era of fully autonomous mining, geotechnical hazard detection systems should also transition to working autonomously to detect hazards in real-time without any human interruption. Subsets of AI, such as deep learning, are likely to be crucial to developing these systems but developing AI-based hazard detection systems has its unique challenges. Firstly, deep learning development requires many examples of different output classes for training. This requirement comes at a significant cost since developing deep learning systems for geotechnical hazard detection requires on-site data collection which can be time-consuming and can lead to additional expenses for the organizations depending on the tools needed to collect the necessary data. In addition, some areas may not be accessible for data collection due to safety concerns or operational restrictions. These limitations on the data collection result in small and unbalanced datasets that reduce the performance of the deep learning systems. Secondly, deep learning systems operate as black-box algorithms, meaning that the system is not human-understandable and there is no logic to its predictions. The lack of transparency associated with these models decreases the confidence of the users to utilize the systems on critical tasks such as hazard detection since the results of a failure can be catastrophic. On the other hand, an autonomous geotechnical hazard detection model is an essential part of mine autonomy, especially for underground mining operations. Moreover, development of an autonomous geotechnical hazard detection requires integration of various information components, namely, roof fall hazard management systems in various mining operations (e.g. underground coal, stone/hardrock mines, tunneling), applications of deep learning in mining and tunneling studies, various deep learning systems (e.g. deep learning with small data sets, transfer learning etc.), and use of synthetic data in deep learning models. In the following subsections, a literature review on these information components are given to set the motivation, goal and contribution of the thesis work.

1.1 State of the Art for the Essential Components of an Autonomous Geotechnical Hazard Detection System

1.1.1 Roof Fall Hazard Management

Roof falls are defined as the fall of material from roof and walls in underground mining and tunneling operations due to geological and operational conditions. Geological conditions involve the presence and orientation of discontinuities, stresses around the opening, geo-mechanical properties of roof materials, strata thickness, and moisture content. Operational conditions include mining method, locality, height, and width of the openings. Seasonal effects of temperature and humidity also impact roof fall hazards (Pappas and Mark, 2012). Roof falls are responsible for a significant portion of accidents in underground mines, causing fatalities, injuries, damages to equipment, and lost work times (MSHA Part 50, 2019).

Various roof fall hazard management techniques have been implemented by the mining industry. Since the characteristics of roof fall hazards depend very much on the geological and operational conditions, researchers who investigated roof fall hazards previously, chose to limit their studies to a single mining method or a single mine. In this thesis, the same tradition is followed with an aim of generalizing across different mines in the future.

1.1.1.1 Roof Fall Hazard Management in Coal Mines

Duzgun and Einstein (2004) proposed a risk analysis methodology for the management of roof fall risks in coal mines using past roof fall accident statistics. This methodology included the probability of a roof fall, possible consequences, and the cost of consequences. The decision framework developed in this study proposes alternative actions depending on the level of risk ranging from “do nothing” to “support improvement”. Duzgun (2005) expanded the proposed decision framework to a different coal basin using different probability distributions to account for the different nature of the roof fall hazards.

Shen et al. (2006) described the development of a roof monitoring system that integrates displacement, and seismic monitoring. The data collected by the system produced indicators of future roof falls in longwall coal mining operations. However, in this methodology, the future roof fall predictions depend on displacement. Therefore, the prediction of the hazard may already be too late to take appropriate action since the displacement already started. For autonomous mining, the timely indication of hazards is crucial to ensure the continuity of the operation without disruptions.

Ghasemi et al. (2012) discussed effective parameters on roof falls during retreat mining. These parameters were drawn from past experiences as well as relevant literature, and they were categorized into three groups: geological, design, and operational parameters. The authors proposed an equation to calculate roof fall risk during retreat mining by quantifying these parameters. Although the methodology is useful in new mining areas where the personnel cannot identify hazards based on experience, it still requires expert interruption to reach high performance and, therefore, may not be as effective in autonomous operations.

Prusek et al. (2017) used expert opinions and empirical methods to develop a model for roof fall risk assessment in longwall mines. Their method calculates the probability of a roof fall based on the roof bearing capacity index and quantification of seven geological, mining, and technical factors obtained through the examination of the literature. This methodology allows mine management to obtain information regarding roof fall hazards. However, the calculation of the roof fall probability requires an on-site inspection from experienced personnel, therefore it is not suitable for autonomous mining operations.

Zhao et al. (2016) proposed a sensor-based roof monitoring system for coal mines that detects pressure changes in roof layers. The methodology provides opportunities for mine management to monitor roof conditions in real-time. However, the correlation between the pressure changes and the roof fall hazards should be investigated for autonomous hazard detection.

Sherizadeh and Kulatilake (2016) used a numerical modeling technique to capture the geo-mechanical behavior of an underground mine with a roof instability problem and incorporated a roof fall detection criterion in their framework. The study identified that the bedding planes play an important role in roof stability and the model was able to accurately capture possible failure of major discontinuities.

Shen et al. (2018) established a numerical model to predict the displacement for the entry roof with a weak plane. The authors used the numerical model to test the effect of different width to height ratios of the coal pillar to reduce the effect of mining stress on roof displacements.

Ghosh and Sivakumar (2018) implemented a real-time, micro-seismic monitoring technique for roof fall hazard management in longwall operations. The parameters obtained using this technique have helped predict roof fall events and understand fracture propagation patterns. However, the results of this technique only provide prior information to mining personnel to assist them in their decision-making regarding roof fall hazards. Therefore, the technique is not suitable for autonomous applications.

To summarize, in coal mines, roof fall hazard management techniques mostly depend on sensors, data analytics on past roof fall cases and the decision-making skills of experienced personnel. Also, some techniques focused on tackling the problem at the design stage using numerical analysis. The methodologies listed here are all developed to provide some form of technical information to experienced personnel to help them evaluate the roof fall hazards. Therefore, the roof fall hazard management techniques developed for coal mining are not suitable for autonomous mining operations.

1.1.1.2 Roof Fall Hazard Management in Stone Mines

Underground stone mines are even more prone to roof fall hazards as the openings are much taller compared to coal mines. Iannacchione et al. (2005) investigated the use of micro-seismic monitoring to predict roof falls in stone mining. 90% of the predicted roof falls had a warning time greater than one minute. Although this warning time was promising, about 50% of the roof falls forecasted by this method were false alarms. An autonomous hazard detection system needs to have higher accuracy in order not to hurt the productivity improvements gained by the autonomous operations.

Iannacchione et al. (2006) proposed a Roof Fall Risk Index (RFRI) to quantify the roof fall hazards in underground stone mines using an observational technique. The RFRI consists of ten "defect" categories that include geology, mining-induced factors, roof profile, and moisture conditions. These categories were determined based on the extensive experience of the authors with underground stone mines, combined with an examination of the literature. Iannacchione et al. (2007) combined RFRI with a risk assessment technique. In this methodology, RFRI is used to identify potential roof fall hazards, which is used as a proxy for the probability of a roof fall occurring. The other parameters used in the calculation of the risk are the potential of a miner being injured and the severity of the roof fall event. Mine engineers can then display roof fall hazards on a risk map generated by this methodology. These methodologies provide an assessment tool to decision-makers to help them with their engineering judgment. Since the determination of the RFRI depends on visual observation it is not a suitable technique for autonomous operations.

Bertoncini and Hinders (2010) used fuzzy classification to predict roof fall events in limestone mines using microseismic data. Bertoncini and Hinders' classifier successfully forecast two major roof fall events more than 15 hours before the first visual signs. However, the authors noted that smaller roof falls may have occurred and gone unrecorded.

In stone mining, the methods developed to mitigate roof fall hazards are mostly based on seismic monitoring and visual observation, and the decision-making skills of experts. However, these methods would not apply to autonomous operations due to their requirement of personnel presence and high error rate. Finally, after analyzing the roof fall hazard management techniques developed for the mining industry, one can argue that the visual observation and decision-making skills of experienced personnel is highly valuable for detecting the hazards.

1.1.1.3 Roof Fall Hazard Management in Tunneling

Qin and Chian (2018) presented a kinematic analysis methodology for roof stability investigations in deep-buried tunnels. The methodology accounts for the non-uniformity of rock properties. The collapse mechanisms provide insight into the implications of influential factors on roof stability in deep tunnels.

Shi et al. (2019) used a support vector machine method to predict deformation in shallow-buried tunnels. The method predicts the minimum, average and maximum deformation of the surrounding rock formations. The predictions agree well with their applied case study.

Sengani (2020) investigated the use of ground-penetrating radar to detect and distinguish seismic and non-seismic hazards in mining/tunneling. The information acquired by the scanner provided useful input for numerical analysis. Geological structures such as faults, joints, and bedding planes were identified. Also, the rib pillars were calibrated based on the scanner results.

These recent studies in roof fall hazard management in tunneling focused on managing roof fall problems at the design stage using numerical modeling of the excavations and advanced sensor technologies. Although managing the hazards at the design stage has various advantages, autonomous operations also require a hazard detection system to ensure the continuity of the operation.

1.1.2 Deep Learning Applications in Mining and Tunneling

Mining and tunneling industries have started using AI for various purposes in the last few years. Huang et al. (2018) proposed a novel image recognition algorithm based on deep learning to recognize crack and leakage defects of metro shield tunnels. The proposed method has better performance in terms of recognition results, inference time, and error rates compared to traditional methods.

Huang et al. (2018) used deep learning to identify the source location of micro-seismic events in underground mines. Experimental data from blast tests show that the proposed approach can identify the event source locations. Identifying the micro-seismic event location can be useful to detect hazards in mining such as rockburst, roof caving, water inrush, and slope landslide.

Han et al. (2019) trained a convolutional neural network (CNN) using spectrograms of sound frequencies created by striking a rock with a geological hammer to measure rock surface strength. The methodology is an objective rock surface strength measurement technique with a 93% classification accuracy.

He et al. (2019) developed a method to estimate field strength parameters of rock continuously by incorporating a CNN into the drilling process. Experimental results showed that the predicted parameters are within the accepted error range of 10%. The method enables reliable field measurement much faster than the standard tests.

Xue et al. (2020) developed a deep learning-based method to automatically and accurately calculate the water leakage area in shield tunnel lining from images. Since the number of water leakage pixels cannot be regarded as reliable metrics to evaluate the safety of the shield tunnel in terms of water leakage, a series of field experiments were conducted to obtain the calibration relationship between the number of target pixels and true areas. The average error rate was 2.6%, which is within the tolerance of the engineering.

Zhao et al. (2020) presented an image segmentation method for moisture marks of shield tunnel lining using CNN. The proposed method is validated by an experimental study, and the results are compared with those obtained by conventional methods. The accuracy of the proposed method is higher compared to conventional methods.

Wilkins et al. (2020) constructed CNN to identify micro-seismic events. Locating micro-seismic events can potentially forecast catastrophic failures. However, small micro-seismic events and mining activities create noise.

Wu et al. (2020) developed an AI system to monitor tunnel construction activities by integrating domain expertise to improve system accuracy. The system uses an AI framework that identifies construction works using low-cost site surveillance images. The framework estimates the degree of plan-work deviation and identifies work categories inside each work cycle.

Chen et al. (2021) created a database of tunnel face images from highway tunnel projects and developed a CNN framework for classification of multiple rock structures. The authors compared different CNN architectures and different hyperparameters to optimize the framework.

The number of deep learning applications in mining and tunneling has been increasing rapidly. However, it has not been used for geotechnical hazard detection. The promising results of these studies present deep learning as a strong alternative for developing hazard detection systems.

1.1.3 Deep Learning Applications with Small Datasets

The most significant challenge of developing an AI-based geotechnical hazard detection system is that it requires a large number of examples from which to learn. In the case of roof fall hazard detection, data collection needs to be done on-site which presents more difficulties. On-site data collection may limit the number of data points since some areas may not be accessible for data collection due to safety concerns and operational restrictions. Various techniques have been proposed in the literature to deal with the problem of small and unbalanced datasets in deep learning. The most commonly used techniques are transfer learning, synthetic data generation, and data sampling. The applications of these techniques are explained in the following subsections.

1.1.3.1 Transfer Learning Applications

Transfer learning has emerged as an alternative when insufficient data prevents achieving acceptable performance on classification tasks. In transfer learning, a network trained in one domain of interest is used for classification in another domain of interest with fewer training data (Pan and Yang, 2010). The idea behind transfer learning is that many networks learn low-level features that are not specific to any one image class and can, therefore, be used in other tasks. Practically, this means we can take an existing model to recognize novel image classes by retraining only the last layers where features are aggregated into objects. The transferability of features decreases as the similarity between the original task and target task decreases. However, Yosinski et al. (2014) showed that transferring features even from distant tasks can give better results than using random features.

An example of using transfer learning in geosciences is presented by Li et al. (2017). They applied the technique to the classification of sandstone images. The results proved the effectiveness and the validity of transfer learning on increasing the performance of microscopic interregional sandstone image classification.

Cunha et al. (2020) used transfer learning to train an existing CNN classifier to detect faults based on seismic data. The classifier was pre-trained with synthetic data and it is used in the classification of real data. The final layer of CNN was trained with 756 real images in a very short time and the classifier gave satisfying results.

Li et al. (2020) used a CNN classifier developed with a transfer learning approach to automatically classify four groups of Martian rocks. The proposed model has the best performance on Martian rock classification when compared to other models. The number of images used in this study is 620.

Fan et al. (2020) developed a CNN classifier with transfer learning for the lithological classification of 28 rock types. The classification accuracy and time meet the requirements of fieldwork. 2500 images were used to train the final layer of the pre-trained network.

These studies show that transfer learning is an effective approach to train a classifier with small datasets in geosciences. Therefore, its application in geotechnical hazard detection is also promising.

1.1.3.2 Synthetic Data Generation Applications

Another solution to the small dataset problem in deep learning is to use synthetic data. Synthetic data can be defined as the data artificially manufactured by a computer instead of real-world measurements. In most cases, manufacturing synthetic data is more cost-effective and timesaving than collecting real-world data. Also, synthetic data can be used to replace, or augment limited real-world data so that the training dataset includes examples of every imaginable scenario.

Planche et al. (2017) generated realistic depth images from CAD models for training in recognition tasks. The authors were able to comprehensively model all vital factors such as noise, reflectance, and surface geometry. The integration of synthetic data in this recognition framework enhanced the performance significantly.

Wan et al. (2017) proposed a variational autoencoder-based synthetic data generation method which generates new samples that are like the original data but not the same. The authors used synthetic images to reduce class imbalances. The comparison of the proposed method with other sampling methods developed for class unbalance demonstrated the effectiveness of it.

Tremblay et al. (2018) presented a system for training deep neural networks for object detection using synthetic images. The synthetic images are rendered from 3-D models of cars. The other environment variables such as lighting, camera angle, and texture are randomized for each render. The authors have shown that this approach outperforms networks trained with real data alone.

Ward et al. (2018) used a combination of real and synthetic plant images for leaf segmentation and the proposed approach outperformed existing methods in the field. Synthetic images are generated by using variations in the geometry and sampling of real-world textures for background and foreground.

Luo et al. (2019) proposed a synthetic data generation method to deal with the extremely unbalanced data problem in anomaly detection. Synthetic abnormal data samples are generated by a machine learning method. The authors showed the method performs steadily in multiple anomaly detection fields such as fraud and disease detection.

These studies show that replacing or augmenting deep learning training sets with synthetic images is an effective approach to deal with the problem of small datasets. Therefore, using synthetic images when developing an AI-based geotechnical hazard detection system also seems promising.

1.1.3.3 Data Sampling Applications

Data sampling can be used as a solution to the class imbalance problem in which one class is underrepresented in the training dataset. In data sampling, a balanced dataset can be achieved by duplicating and augmenting the original data in the minority class (oversampling), or by randomly removing data from the majority class (under-sampling). Although these techniques are easy to implement, they may bring some problems such as overfitting and losing important information. Johnson and Khoshgoftaar (2020) discussed the effects of different data sampling techniques on deep learning. Also, more complex sampling methods have been proposed in the literature.

Van Grinsven et al. (2016) proposed a selective sampling method that samples training data based on the current CNN status. The method assigns weights to training samples and the informative samples are more likely to be included in the next training iteration. This method increased the classification performance on the detection of hemorrhages while decreasing the training time.

Douzas and Bacao (2017) proposed a sampling method using self-organizing maps. The method produces a two-dimensional representation of the original high-dimensional space of the unbalanced dataset. Then, within-cluster and between cluster synthetic samples are generated.

Ando and Huang (2017) proposed a framework for extending the oversampling method to the deep feature space acquired by CNN. The framework extends the synthetic over-sampling technique not only to complement the minority classes but also to improve dataset quality enhancing class distinction. The results showed that the proposed framework can address the class imbalance more effectively than the existing methods.

These studies show that data sampling techniques that improve dataset quality by assigning weights to the informative samples or eliminating non-informative samples are effective to improve classification performance. Therefore, implementing a data sampling technique when developing an AI-based geotechnical hazard detection system seems promising.

1.2 The Research Goal

The use of deep learning in geotechnical hazard management has not been reported. This thesis aims to investigate the use of deep learning in developing autonomous geotechnical hazard management systems. For this purpose, a case study mine that experiences frequent roof falls is chosen. An AI-based roof fall hazard detection system is developed using limited image data and the methods to improve deep learning performance are investigated. Also, to create an interpretable system, the relationship between the model's predictions and its input features are examined.

1.3 Research Questions

The research questions of this thesis and the associated reasons to explore the answers to these questions are listed below:

1) Is transfer learning effective in training an AI-based roof fall hazard detection system with small datasets?

On-site data collection required for developing deep learning applications for geotechnical analysis may present several restrictions that prevent engineers from collecting large amounts of data. Applications of transfer learning with small datasets presented in the literature showed that it is possible to transfer features from a task trained with a large dataset for classification in another task with small dataset. Therefore, utilizing transfer learning to develop

AI-based roof fall hazard detection systems may significantly reduce the amount of data required.

2) What is the effect of replacing or augmenting real data with synthetic data on the performance of AI-based roof fall hazard detection systems?

The use of synthetic data is useful in developing deep learning models in various domains. In the case of hazard detection, replacing or augmenting real data with synthetic data may reduce the personnel exposure to safety hazards during data collection. Therefore, adopting a synthetic data generation step within the methodology has the potential to improve the classification performance by expanding the training set with new data which would not be possible to collect otherwise.

3) What is the effect of data sampling on the performance of AI-based roof fall hazard detection systems?

Data sampling techniques have been shown to improve the success of deep learning models by solving class unbalance problem or by removing non-informative samples. Utilizing a data sampling technique to improve the quality of the dataset by removing non-informative samples may improve the performance of the AI-based roof fall hazard detection systems.

4) What are the similarities between the features used by an expert and AI-based roof fall hazard detection system to predict roof fall hazards?

Literature review on roof fall hazard management revealed that expert opinion is highly valuable and is still commonly used as part of the hazard management systems. In such cases, an AI-based roof fall hazard detection system should utilize the same features as the expert for hazard prediction. Therefore, the predictions made by the AI-based roof fall hazard detection system should be interpretable, to build the confidence of the users.

1.4 The Original Contribution of the Research

Deep learning has not been used in the automated visual detection of geotechnical hazards despite its usefulness in other fields. In this thesis, the limitations of deep learning that caused researchers to not use it as a hazard detection tool, and the methods to overcome these limitations are presented. This research provides the first AI-based geotechnical hazard detection system by implementing these methods.

It is known that deep learning requires many examples of different output classes during training. Some researchers attempted to develop deep learning models with limited training sets using a variety of techniques. This thesis presents the first methodology in the domain of mining engineering to develop a deep learning system with small datasets using a combination of techniques, namely transfer learning, synthetic data generation, and data sampling.

Fully autonomous mining operations require autonomy in every component of a mining system including hazard detection. This thesis presents the first autonomous geotechnical hazard detection system that can operate without a human in the loop. Furthermore, the system is the first application of interpretable deep learning in mining/tunneling, which means that it provides explanations in terms of input features for its predictions.

1.5 Structure of the Thesis

This thesis has eight chapters. Chapter Two presents the roof fall problem and the roof fall hazard management techniques implemented in the case study mine. Chapter Three introduces the methodology followed in this research. Chapter Four analyzes the association between the expert decision-making and previous roof fall events in the case study mine, which forms the basis for data collection and labeling for deep learning development. Chapter Five presents the use of transfer learning to develop the AI-based roof fall hazard detection system.

Chapter Six presents the investigation of performance improvements by replacing or augmenting real data with synthetic data and data sampling. Chapter Seven presents the final testing of the AI-based roof fall hazard detection system and the interpretation of the predictions made by the system. Chapter Eight presents the conclusions and recommendations of the dissertation.

CHAPTER 2

DESCRIPTION OF THE CASE STUDY

The Subtropolis mine is an underground limestone mine in eastern Ohio. It has been operational since 2006. The mining method is room and pillar in which rooms are extracted and pillars are left in place for support (Figure 2.1).

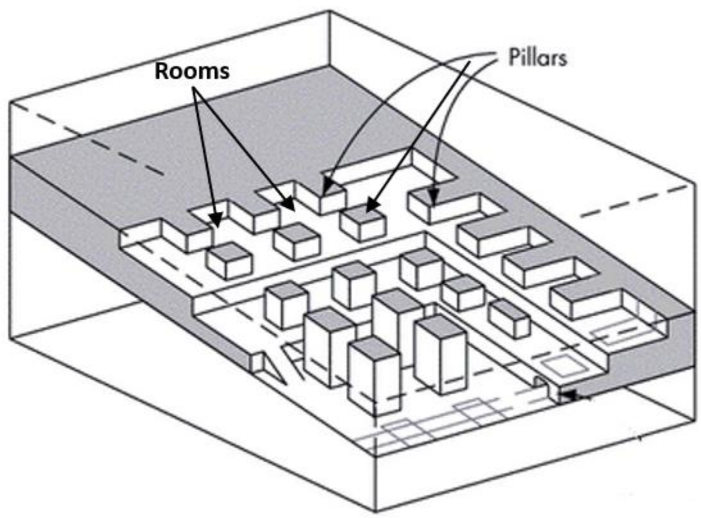


Figure 2.1 Room and pillar mining method (Source: SME Mining Engineering Handbook, by P. Darling, 2011).

Figure 2.1 shows a typical room and pillar mining operation. In Subtropolis, rooms are 9 – 12 m wide and 5 m high. The rectangular pillars have a length-to-width ratio ranging from 1.5 to 2, with lengths of 14 m and widths of 8 m. The Subtropolis mine has a history of frequent roof falls originated from high horizontal stresses. Figure 2.2 shows the aftermath of a roof fall in the Subtropolis Mine.

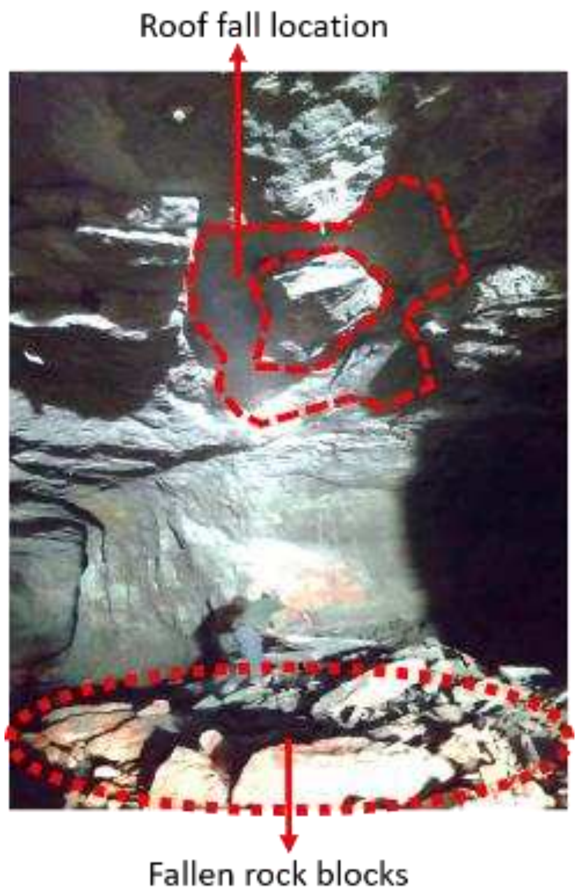


Figure 2.2 Roof fall in Subtropolis Mine

2.1 Stress Conditions Around the Region

Horizontal stresses cause severe ground control problems in underground limestone and coal mines throughout the Eastern and Midwestern United States (Iannacchione et al., 2002; Mark and Mucho, 1994). The concentration of horizontal stresses in bedded deposits originates from plate tectonics. Stress measurements around the Eastern and Midwestern United States show that the direction of maximum horizontal compressive stress ranges from N 70°E to N 80°E in most mines. This direction is the same as the direction of movement as the North American Plate moves away from the Mid-Atlantic Ridge at a rate of ~2.5 cm/year (Figure 2.3) (Iannacchione et al., 2002).

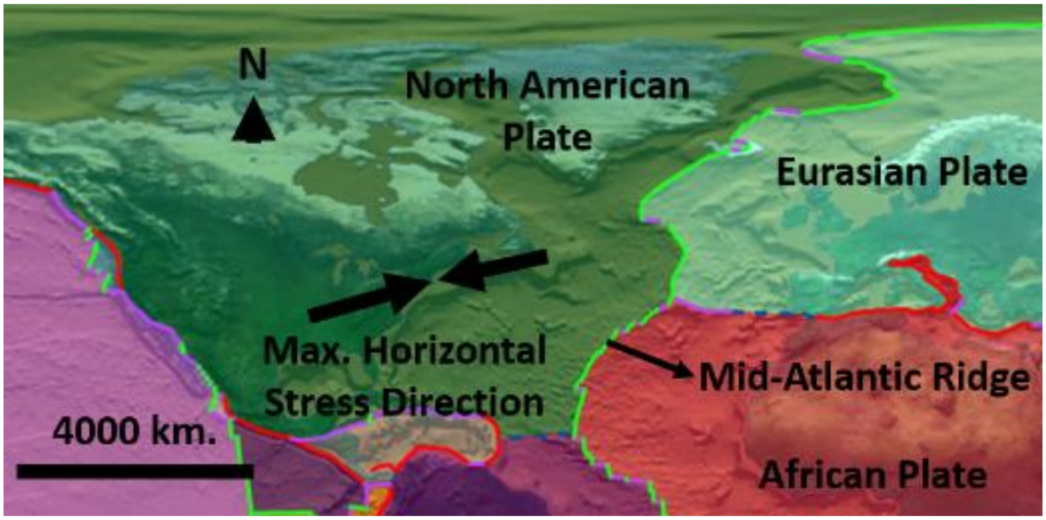


Figure 2.3 Effect of plate tectonics on horizontal stress on bedded deposits around Eastern and Midwestern United States (Adapted from NASA/Goddard Space Flight Center Scientific Visualization Studio, 2004)

The tectonic stresses around the region build a constant strain field (between 0.00045 – 0.00090 Ɛ), which induces higher levels of horizontal stress in soft limestone formations. This explains why some underground stone mines experience high levels of horizontal stress (Esterhuizen et al., 2008). Dolinar (2003) also shows that variations in the magnitude of the horizontal stresses in the Eastern and Midwestern United States are better explained by the elastic modulus of the rock, than by overburden depth.

2.2 Stress Conditions at the Subtropolis Mine

In Subtropolis, the mine layout has been developed to control high horizontal stress concentrations, with varying degrees of success. The advantage of adopting a stress control layout is that it maximizes the number of headings driven parallel to the direction of maximum stress, thereby reducing stress-related ground control problems. The Subtropolis Mine implemented a new stress control layout after experiencing frequent roof fall problems. Iannacchione et al. (2020) provided a thorough explanation of the mine layout considerations in the Subtropolis Mine. The headings advanced in East-West direction until 2007 when the orientation of the stress field was thought to be closer to North-South, and the Subtropolis mine layout was reoriented. A North-South mining orientation was then used until 2018 when the ground conditions worsened, and mining operations had to be paused in several faces. Horizontal stresses induce large oval-shaped roof falls, and the long-axis orientation of the ovals is generally at right angles to the direction of maximum horizontal compressive stress. In the Subtropolis mine, this orientation was found to be around N 55°E. Therefore, the direction of horizontal stress should be N 35°W, which was confirmed by an investigation of strata within roof exploration boreholes. The maximum horizontal stress direction was found to be the same in two other underground limestone mines in the same area. However, this direction is different from the direction of maximum horizontal compressive stress in Eastern and Midwestern U.S. which is usually between N 70°E to N 80°E. The reason behind the difference has not been reported in the literature, and the mine management did not investigate it.

Following the correct determination of maximum horizontal stress direction, all new headings have been aligned N 35°W. With this new mine layout, ground control issues have been transferred to the crosscuts, which mine management expected.

2.3 Roof Fall Hazard Management at the Subtropolis Mine

Roof fall hazards are usually mitigated by various methods such as scaling of the weak roof layer, roof bolting, stress control mine layouts, seismic monitoring, and convergence detection with laser scanning. In Subtropolis mine, three roof fall hazard mitigation and detection methods are implemented. The first method is the stress control mine layout which was explained in Section 2.2. The second method is roof bolting and scaling. The third method is daily routine inspections by ground control personnel. The last roof fall hazard mitigation method will be the main focus of this research for the remainder of the thesis.

Ground control personnel at the Subtropolis Mine makes routine daily inspections to detect and keep track of roof fall hazards in the mine. Through years of experience in the same geological and operational domain, they have been able to detect roof fall hazardous areas successfully. High horizontal stresses produce visual clues in the form of roof beams since the roof beams subjected to horizontal stresses deform or sag more than those only subjected to gravity (Iannacchione et al., 1998). The ground control personnel at Subtropolis mine have been able to associate areas of high horizontal stress with roof beams. Some examples of stress-induced roof beams are shown in Figure 2.4.

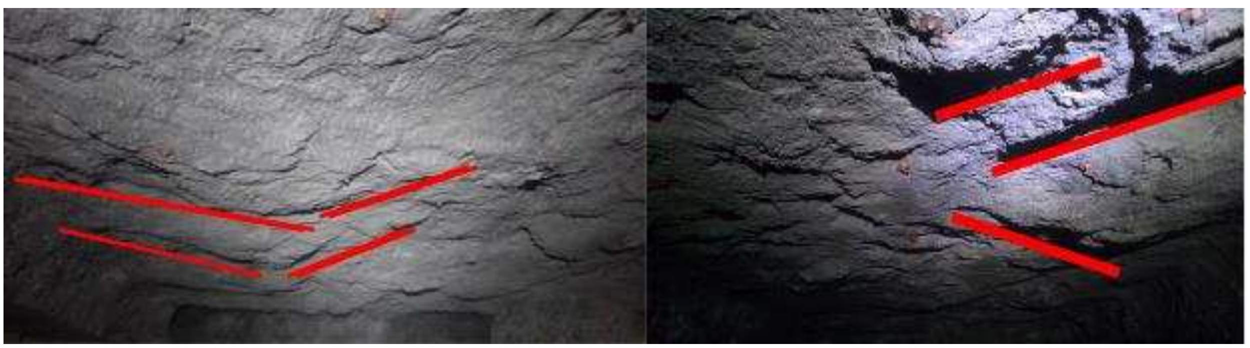


Figure 2.4 Stress-induced roof beams

Roof beams are used as indicators of hazardous roof conditions. The ground control personnel identified the frequency and the depth of roof beams as the most important characteristics of the roof beams in defining hazard levels. Based on ground control personnel's hazard detection using roof beams, mine management takes safety precautions such as roof support improvement and restricting access to hazardous areas. To keep track of the roof fall hazards in the mine and improve hazard management, ground control personnel regularly map the roof beams. Figure 2.5 shows a section of the roof fall hazard map created by the ground control personnel.

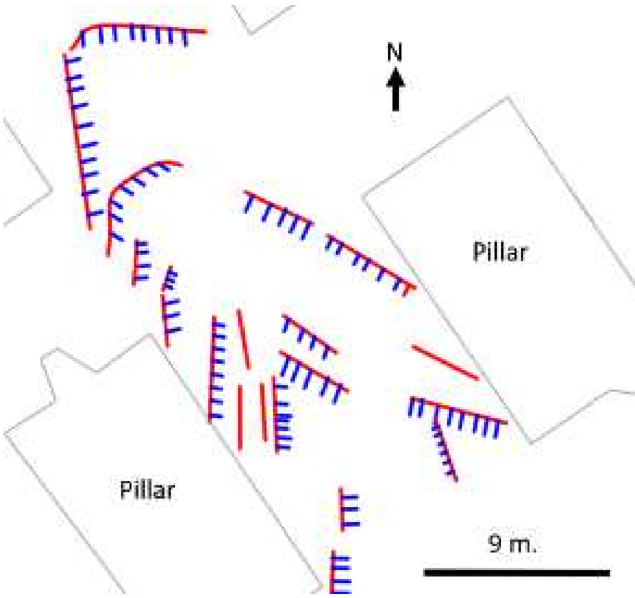


Figure 2.5 Roof beams on the mine roof fall hazard map

Each red line on the map represents a roof beam. Blue lines drawn perpendicular to the red lines represent the depth of a roof beam. The map is created and updated regularly by the ground control personnel based on their visual interpretation. The frequency and the depth of roof beams can be interpreted using this map which enables the personnel to keep track of the hazardous areas and implement safety measures.

CHAPTER 3

METHODOLOGY

In this chapter, the methodology followed in this thesis is explained. Stress-induced roof beams have a crucial role in detecting roof fall hazards in Subtropolis mine. Ground control personnel use visual observation to associate the depth and the frequency of the roof beams to roof fall hazard levels. Since roof fall hazard detection depends on visual observations, we sought a system to work with similar visual inputs. Therefore, a (CNN) trained to recognize a variety of visual features that the rest of the model can use for classification or object detection is chosen. This task requires collecting images of hazardous and non-hazardous roof conditions. However, hazardous roof areas are few compared to non-hazardous roof areas. Besides, collecting images in hazardous roof areas may not always be possible due to safety concerns and operational restrictions. For these reasons, it is expected that the roof image dataset is limited and unbalanced. Therefore, synthetic data generation and data sampling techniques are adopted to train CNN with a transfer learning approach. To understand the input features used by the model for hazard prediction, a model interpretability algorithm is utilized. Figure 3.1 (page 25) illustrates the steps of the research methodology.

In Step One, it is shown that the depth and frequency of roof beams correlate with the locations where roof falls occurred previously in the mine. This verifies the accuracy of the ground control personnel's description of their decision-making skills regarding the detection of roof fall hazards, with quantitative methods. The result of this analysis is used to justify the labeling of the data collected.

In Step Two, first, the collection of image data used in developing the AI-based roof-fall-hazard detection system is explained. The image data is collected under two classes, namely hazardous roof conditions and non-hazardous roof conditions. The data collection locations are selected based on the recommendation of the ground control personnel. Also, using quantitative methods, it is verified that the hazardous and non-hazardous data collection locations exhibit different roof fall hazard levels. Second, the images are pre-processed to increase the size of the image dataset and resized as required by the CNN algorithm. Finally, the selected CNN architecture is explained, and the network is trained using a transfer learning approach. In this approach, a network pre-trained on 1,200,000 images is utilized as a starting point to train the final fully connected layer of the network using the hazardous and non-hazardous roof images.

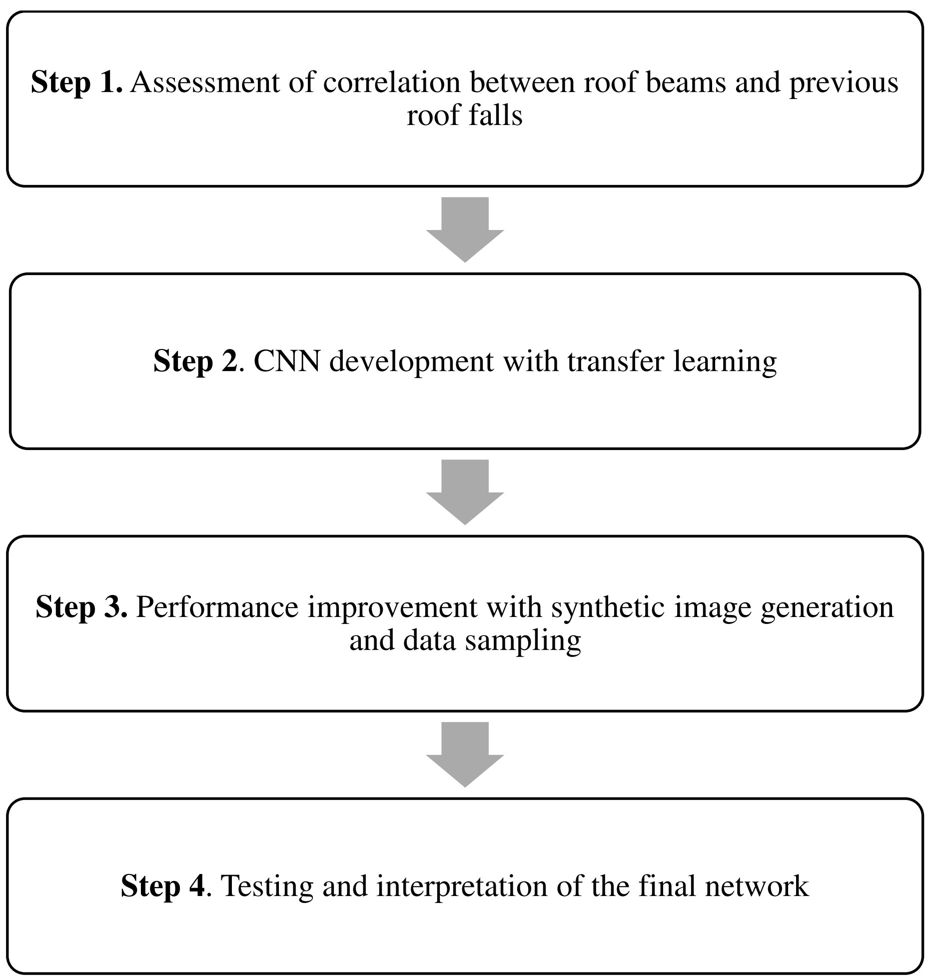


Figure 3.1 Research methodology

In Step Three, first, a 3-D model of a hazardous roof condition is generated using close-range digital photogrammetry. Synthetic images are rendered from the 3-D model. The rendered images are used to replace or augment real data. New training sets are created using real and synthetic images, and a network is trained with each new training sets. Network performances are compared with that of the previous step. Second, a simple data sampling technique based on Gabor magnitudes is introduced. The data sampling technique aims to improve the quality of the training set by removing non-informative images from the dataset. New training sets are created using different data sampling levels, and a network is trained with each new training sets. Network performances are compared with that of the previous step.

In Step Four, first, the performance of the final network is verified using the test dataset. The test images are taken at the mine entries that were not used in the previous data collection steps. Second, a deep learning interpretation technique is utilized the provide explanations on the predictions made by the classifier. The technique shows the most important pixels for the prediction. The details of each stage are given in the following chapters.

CHAPTER 4

ANALYSIS OF CORRELATION BETWEEN ROOF BEAMS AND PREVIOUS ROOF FALLS

Submitted to the Tunnelling and Underground Space Technology Journal

Ergin Isleyen[*1], Sebnem H. Duzgun[1], McKell R. Carter[2]

Ground control personnel at the Subtropolis Mine use their experience-based skills from years of mining experience in the region to detect roof fall hazards from stress-induced roof beams. They have identified the depth and the frequency as the most important characteristics of the roof beams. Therefore, they map the roof beams in a way that represents depth and frequency. In this chapter, the correlation between the characteristics of roof beams and previous roof falls is demonstrated using statistical analysis.

4.1 Data Collection

In this chapter, the mine map created by the ground control personnel to keep track of the roof fall hazards is used for analyzing the correlation between roof beams and previous roof fall locations. The map includes the location and the depth of roof beams. This map was acquired by the mine management on October 8th, 2019.

Figure 4.1 shows the location of 58 known roof falls provided by ground control personnel. These were compared to 58 randomly selected locations with no previous roof falls.

[*] Corresponding author. Direct correspondence to isleyen@mines.edu
1 Mining Engineering Department, Colorado School of Mines, 1500 Illinois St., Golden, CO 80401, USA
2 Institute of Cognitive Science, University of Colorado Boulder, Muenzinger D244 345 UCB, Boulder, CO 80309, USA

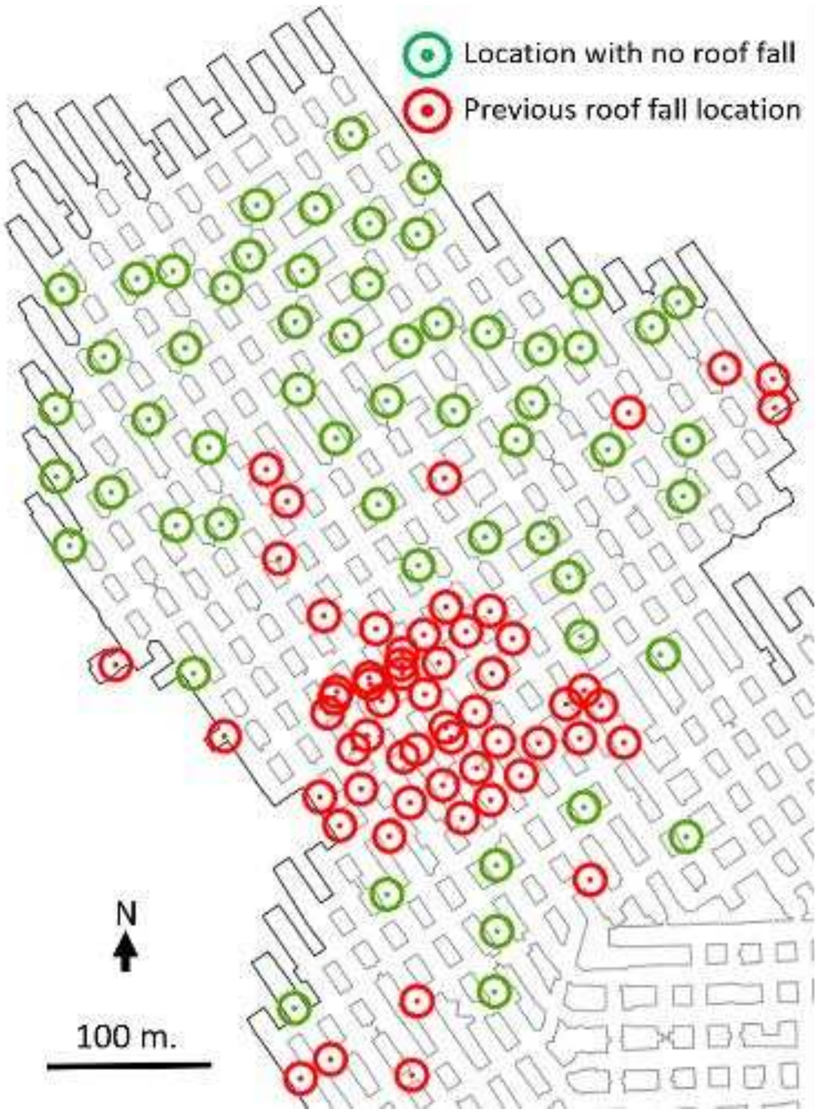


Figure 4.1 Previous roof fall and no-roof fall locations

4.2 Quantification of the Characteristics of Roof Beams

To test for an association between stress-induced roof beams and previous roof falls, a circle with a radius of nine meters was drawn at each of the 116 locations. The area covered by the circles match the size of the rooms, and it is assumed that this area is the effective area of roof beams regarding roof falls by the recommendation of the ground control personnel. At previous roof fall locations, the roof beams mapped before the failure, are used. For each location, the number of roof beams inside each circle is recorded as frequency, and the average depth of the roof beams within each circle is recorded as depth. Figure 4.2 shows an example of the depth and frequency of roof beams inside a circle centered on a previous roof-fall location.

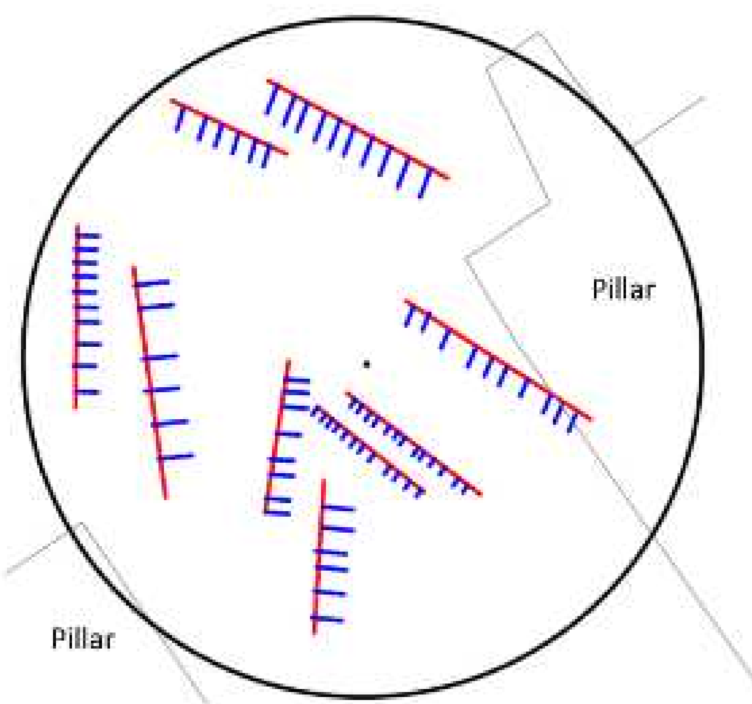


Figure 4.2 Mapping of roof beams around a previous roof fall location

In Figure 4.2, the frequency of roof beams is nine (number of red lines which represent roof beams), and the average depth of roof beams is 2.2 map units (the average length of blue lines which represent the depth of roof beams). The term “map unit” represents the length on the map and not the actual length measurement. Using this technique, frequency and depth information for each previous roof fall location and randomly selected no-roof fall location is recorded.

4.3 Statistical Analysis

The frequency and depth of roof beams between previous roof fall and no-roof fall locations are compared statistically. A statistical t-test is used to test for a difference between the means of the two groups for each feature. Summary statistics and p-values are given in Table 4.1.

Table 4.1 Summary statistics (Mean, Standard Deviation (SD), Degrees of Freedom (df), t-values (t) and P-Values (P-val)

	Frequency		Depth	
	Roof fall	No - roof fall	Roof fall	No - roof fall
Mean	9.76	7.05	1.47	1.17
SD	3.76	3.21	0.65	0.73
df	57		57	
t	4.07		2.28	
P-val	<.001		.02	

For both frequency and depth values of roof beams, the P-values between “previous roof fall” and “no-roof fall” locations are smaller than the level of significance (0.05). The t-test concludes that the difference that was found, or an even larger difference, would occur if the roof conditions were the same (null hypothesis) is less than the level of significance. Therefore, it is plausible that the frequency and depth of roof beams can be used as indicators of a roof fall hazard.

This analysis verifies the success of ground control personnel’s description of their intuitive decision-making skills in identifying features associated with roof falls. It means that the ground control personnel’s visual observation is effective in monitoring the roof fall hazard level. Therefore, it seems promising to target the identification of the features used by ground control personnel as an image recognition problem.

4.4 Results and Discussion

Ground control personnel’s visual interpretation of stress-induced roof beams is used as part of roof fall hazard management at the Subtropolis mine. The depth and frequency of the roof beams are utilized as indicators of a roof fall hazard by the personnel who makes routine inspections to create a hazard map of the mine. Quantitative methods show a strong relationship between roof beams and roof fall incidents. This supports the experts’ description of their intuitive decision-making skills in detecting features that show signs of hazardous roof conditions. Therefore, t-test results indicate that it is plausible to use expert-categorized hazardous and non-hazardous roof locations as a basis for image collection for the CNN image classifier which targets the detection of expert identified features.

CHAPTER 5

CNN DEVELOPMENT WITH TRANSFER LEARNING

Submitted to the Tunnelling and Underground Space Technology Journal

Ergin Isleyen†1, Sebnem H. Duzgun[1], McKell R. Carter[2]

In this chapter, the training of a CNN to detect hazardous roof conditions in Subtropolis Mine is explained. Stress-induced roof beams have a crucial role in detecting roof fall hazards in Subtropolis mine. Ground control personnel uses visual observation to associate the depth and the frequency of the roof beams to roof fall hazard levels. Since the roof fall hazard detection depends on visual observations, a system that works with similar visual inputs is sought. Therefore, a CNN trained to recognize a variety of visual features that the rest of the model can use for classification is chosen. The network is aimed to mimic the visual observation skills of the ground control personnel to detect stress-induced roof beams for the classification of roof images as hazardous or non-hazardous.

The number of images collected in this step is insufficient to train a network from scratch. Therefore, transfer learning approach is adopted to train only the final layer of a pre-trained CNN. The objective of this chapter is to show that the pre-trained network provides the necessary low-level features to train a roof fall hazard detection system using a small dataset.

† Corresponding author. Direct correspondence to isleyen@mines.edu
1 Mining Engineering Department, Colorado School of Mines, 1500 Illinois St., Golden, CO 80401, USA
2 Institute of Cognitive Science, University of Colorado Boulder, Muenzinger D244 345 UCB, Boulder, CO 80309, USA

5.1 Data Collection

Images that will be used to train and validate the CNN are collected under two roof condition classes; hazardous and non-hazardous. The locations where the images are collected were determined based on the recommendations of the ground control personnel at the mine site (Figure 5.1). The images are collected on October 28-30, 2019.

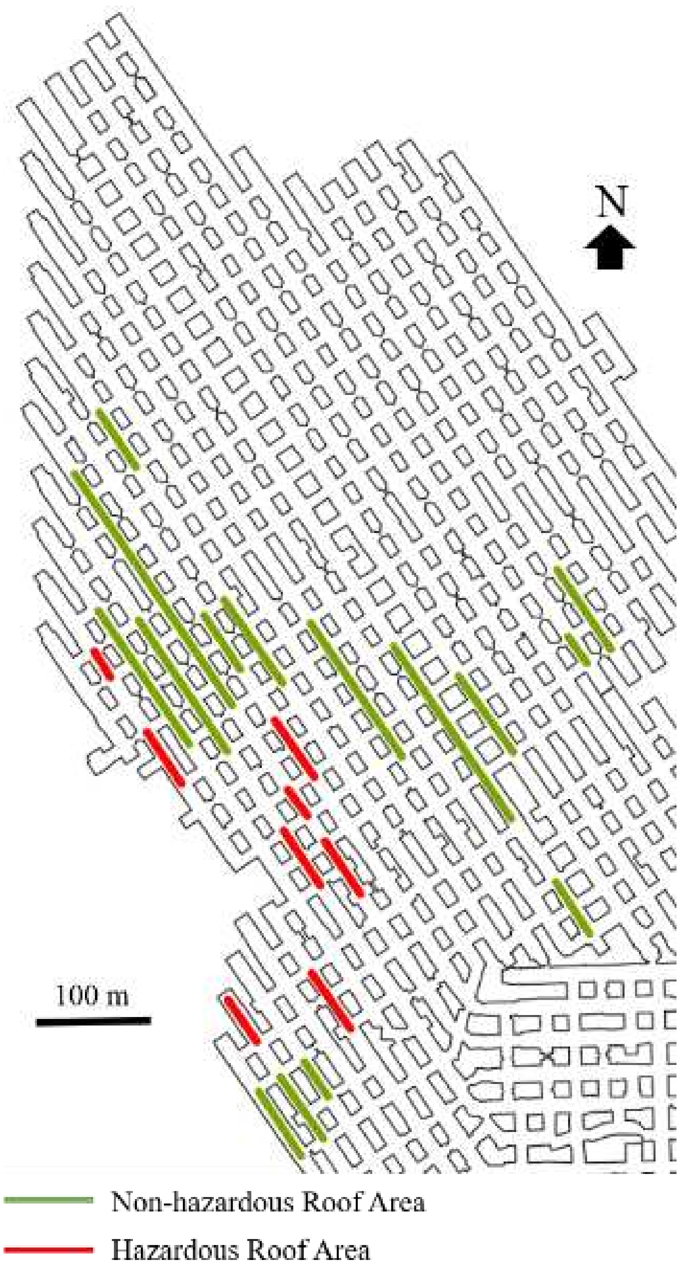


Figure 5.1 Data collection locations

The hazard level difference between locations recommended by the ground control personnel is demonstrated by comparing the means of frequency and depth of roof beams between the two groups with a statistical t-test. The frequency and the average depth of roof beams are measured at 18 m long intervals that are marked for data collection. The measurement of frequency and depth has already been explained in Section 4.1. The results of the statistical analysis for the locations chosen are given in Table 5.1.

Table 5.1 Summary statistics (Mean, Standard Deviation (SD), Degrees of Freedom (df), t-values (t) and P-Values (P-val) for data collection locations

	Frequency		Depth	
	Hazardous	Non-hazardous	Hazardous	Non-hazardous
Mean	8.0	5.9	1.3	0.8
SD	2.4	2.5	0.7	0.6
df	51		51	
t	4.5		3.7	
P-val	<.001		<.001	

The p-values for frequency and depth are smaller than the level of significance (0.05). The t-test concludes that the difference that was found between roof beams, or an even larger difference, would occur if the roof conditions around expert-categorized hazardous and non-hazardous data collection locations were the same (null hypothesis) is less than the level of significance. It also makes it more likely that an image classifier could be used to index roof-fall hazard.

Images were obtained using a Nikon D5000 camera, from locations labeled by ground control personnel as hazardous and non-hazardous. To provide steady lighting during data collection, an LED light source was used. Figure 5.2 shows examples of hazardous and non-hazardous roof condition images. In this step, 83 hazardous and 602 non-hazardous roof images are collected to be used in the training and validation of the CNN.

Figure 5.2 Example images; Hazardous roof (left), Non-hazardous roof (right)

5.2 Image Pre-processing

The number of images collected in the previous step is insufficient to train and validate a CNN. To increase the number of available training and validation images, we first split the images into four tiles (Figure 5.3).

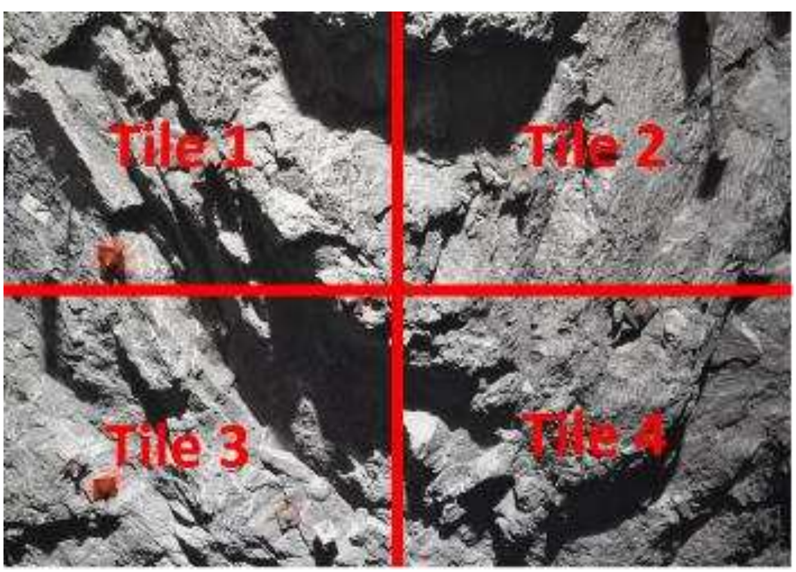


Figure 5.3 Image tiles

The dimensions of the original images are 4288 x 2848 pixels. After tiling, image dimensions reduce to 1072 x 712 pixels. Reduced image dimensions are still large enough to include the roof beams as shown in Figure 5.2. Also, this image size is larger than the standardized input used by the CNN selected for this analysis. Therefore, the images are resized to 224 x 224 pixels as required by the CNN.

In addition to the insufficient number of images, the collected data is highly unbalanced between the two classes. Increasing the number of hazardous images is not possible due to the inaccessibility of other hazardous roof condition areas. To reduce the unbalance, it is decided to maintain hazardous : non-hazardous class ratio of one:two in the datasets. For this purpose, non-hazardous images are randomly removed from the dataset to reach the ratio. A validation dataset is created by separating 20% of the images for hazardous and non-hazardous classes. The rest of the images are used in training. The training set is named training set “A” to distinguish from the other training sets that will be created in the next chapter. Table 5.2 shows the number of images.

Table 5.2 Number of images

Dataset	Hazardous	Non-hazardous	Total
Training (Network A)	266	532	798
Validation	66	132	198
Total	332	664	

Training images goes through a data augmentation stage to artificially increase the number of images seen by the network. The data augmentation method applied in this study rotates each image between 0° - 15°, randomly flips images horizontally, and randomly changes brightness, contrast, and saturation of the images. The data augmentation was only applied to training images, and not on validation images. In this way, the network processes different versions of the same images during the training and validates its accuracy on unaltered versions of the images. Different random transformations were applied at each epoch during the training.

5.3 Development of a CNN

Deep learning is the collective name for computational methods that transform raw input into a representation at a more abstract level with multiple layers of representation. For classification tasks, the input weights from layer to layer are learned to maximize the number of correctly classified images. CNNs are a particular deep learning architecture designed to learn to recognize recurring features. They capably process data that is in the form of multiple arrays, e.g. an image with three color channels. A CNN network typically represents small and meaningful, low-level features of an image, e.g. edges and dots, in the early layers. The later layers recognize objects as combinations of the low-level features (LeCun et al, 2015).

Training a CNN requires a large number of images. In this study, the number of images collected was small even after tiling the images to quadruple the number, and insufficient to train a network from scratch. Therefore, in this research, the transfer learning approach is used to train the network.

Some CNN architectures focus on increasing network depth to improve classification accuracy. However, research has shown that deeper networks are harder to train, and network performance starts to degrade because of the vanishing gradient problem (Glorot and Bengio, 2010). He et al. (2015) addressed this degradation problem by introducing ResNet architecture. In ResNet, layers are reformulated as learning residual functions regarding the layer inputs, instead of learning unreferenced functions. He et al., (2015) presented evidence that these networks are easier to optimize and can gain accuracy with increased depth. In this study, a ResNet CNN architecture is used to develop the AI-based roof fall hazard detection system.

This research performed transfer learning on a ResNet network with 152 layers pre-trained on the ImageNet dataset. ImageNet is an image database that contains 1.2 million images with 1000 categories (e.g. animals, foods, plants, person, etc.), and it is commonly used in object recognition research (Deng et al., 2009). The final fully connected layer of the network is reset and trained with the images collected for this study[‡]. The overall architecture of the network is given in Table 5.3.

Table 5.3 CNN architecture

Output size	152-layer ResNet	Trained on
112x112	7x7 Conv, 64, stride 2	ImageNet
56x56	3x3 Max Pool, stride 2	
	$\begin{bmatrix} 1x1,\ 64 \\ 3x3,\ 64 \\ 1x1,\ 256 \end{bmatrix} x\,3$	
28x28	$\begin{bmatrix} 1x1,\ 128 \\ 3x3,\ 128 \\ 1x1,\ 512 \end{bmatrix} x\,8$	
14x14	$\begin{bmatrix} 1x1,\ 256 \\ 3x3,\ 256 \\ 1x1,\ 1024 \end{bmatrix} x\,36$	
7x7	$\begin{bmatrix} 1x1,\ 512 \\ 3x3,\ 512 \\ 1x1,\ 2048 \end{bmatrix} x\,3$	
1x1	Average Pooling	
	Fully Connected, 2	Subtropolis Roof Images

‡ The source code is available on GitHub at https://github.com/erginisleyen/Isleyen-PhD-Research.

The CNN architecture has 8 stages. Every stage except the fully connected layer was pre-trained on the ImageNet dataset. The network input size is 224 x 224 x 3. In the initial convolution layer, the kernel size is 7 x 7, and the kernel depth is 64 meaning that 64 different kernel filters are used. A stride of 2 means that the kernel slides through the image by moving 2 pixels per step. As a result of the convolution, the output size is reduced to 112 x 112. The objective of the convolution operation is to extract features from the input image. In the second stage, max-pooling performs a down-sampling operation along the spatial dimension using a 3 x 3 kernel. In this stage, the number of parameters is reduced. In stage 3, 3 residual blocks are containing 3 convolution layers each whose kernel sizes are indicated. Similarly, Stage 4, 5, and 6 have 8, 36, and 3 residual blocks respectively each containing 3 convolution layers. The 3 layers are stacked one over the other. The 1 x 1 convolution layers are responsible for reducing and then restoring the dimensions. The 3×3 layer is left as a bottleneck with smaller input/output dimensions. After the convolution layers, the network has an average pooling layer to downscale the image obtained from the previous layers. Finally, the fully connected layer computes the class scores, resulting in the volume of size [1x1x2], where each of the 2 numbers corresponds to a class score. The fully connected layer learns a function using the high-level features obtained as output from the convolutional layers. The total number of parameters in 152-layered ResNet is approximately 60 million. 500,000 of these parameters belong to the fully connected layer. Therefore, in this research, training is limited to these final 500,000 parameters.

The batch size used in this study is 16. The batch size is the number of images that propagates through the network at each iteration. Smaller batch sizes allow learning to start before the algorithm sees the majority of the data, which leads to faster convergence, whereas larger batch sizes cause significant loss of generalization ability of the model (Keskar et al., 2016). An epoch is when the entire dataset is passed forward and backward through the network once. Iterations are the number of batches needed to complete one epoch.

5.4 Results and Discussion

The results of the CNN image classifier are interpreted using the validation accuracy and the training accuracy (Figure 5.4).

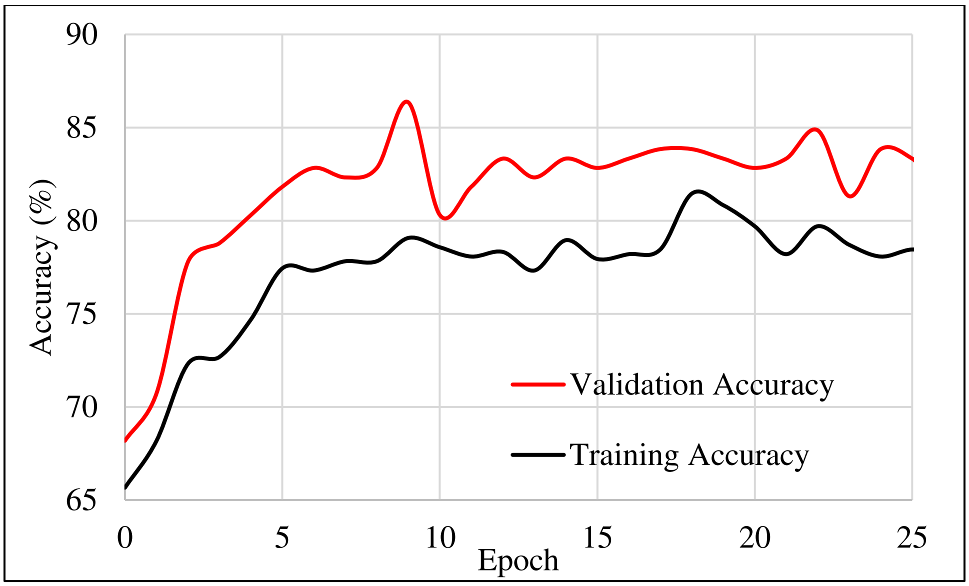


Figure 5.4 Validation and training accuracy

The highest validation accuracy is 86%, reached at the nineth epoch. The out-of-sample validation dataset is used to obtain the validation accuracy of the network. In addition to validation accuracy, the training accuracy is calculated: this is the accuracy of the model when applied to the training images. Training accuracy is used as an indicator of overfitting which means that the model fits too closely with the training data and fails to generalize for the prediction of future data sets. For this network, the general trend of training accuracies is below the validation accuracies. It shows that the model does not overfit to the training data. However, this kind of trend (i.e. training accuracy < validation accuracy) is not commonly observed in

CNNs. A possible explanation for this behavior is the use of dropout layers. Dropout layers reduce overfitting by randomly disabling neurons, which forces the network to work on incomplete representation at subsequent layers (Srivastava et al., 2014). Dropout layers are only used during training and skipped during validation. Therefore, dropout layers are expected to increase training error.

The performance of the model is further analyzed by the confusion matrix that compares the actual and the predicted labels of the validation data (Table 5.4).

Table 5.4 Confusion matrix

		Predicted	
		Hazard	Non-hazard
Actual	Hazard	53	13
	Non-hazard	14	118

The confusion matrix shows that the model successfully predicts hazardous roof conditions with ~80% accuracy (53 out of 66), and successfully predicts non-hazardous roof conditions with ~89% accuracy (118 out of 132). The difference between the accuracy ratio of two classes may be due to the higher number of non-hazardous roof condition training data, resulting in the model's tendency to classify in favor of the non-hazardous class. This results in an increased number of false negatives. In terms of roof fall hazard detection, false negatives create serious safety risks. Therefore, the network performance on hazardous class needs improvement.

The accuracy of the CNN classifier verifies that transfer learning is a useful approach for training a network with an input dataset that is insufficient for training from scratch. The low-level features obtained by the network trained on ImageNet data provide a suitable starting point for training the last layers of the CNN classifier. The accuracy of the classifier shows that these low-level features can be successfully utilized by distant tasks such as detecting roof beams.

CHAPTER 6

PERFORMANCE IMPROVEMENT WITH SYNTHETIC IMAGE GENERATION AND DATA SAMPLING

Submitted to the Geoscience Frontiers Journal

Ergin Isleyen[§1], Sebnem H. Duzgun[1], McKell R. Carter[2]

The performance of the network developed in the previous chapter is limited due to the insufficient number of input images and class unbalance. In this chapter, two techniques that are used for improving the CNN classifier performance are explained, namely synthetic image generation and data sampling.

6.1 Synthetic Image Generation

Data collection for hazardous conditions may not always be possible due to safety concerns. Generating synthetic data using remote sensing techniques is a viable option to create data for hazardous class. Ideally, replacing or augmenting real hazardous data with synthetic data prevents personnel from exposing themselves to hazardous conditions for data collection. In this chapter, a synthetic hazardous roof image dataset is created by rendering images from a 3-D model of a hazardous roof condition developed with close-range digital photogrammetry. The rendered images are used to replace or augment the real data. New training sets are created, and a

---

[§] Corresponding author. Direct correspondence to isleyen@mines.edu
1 Mining Engineering Department, Colorado School of Mines, 1500 Illinois St., Golden, CO 80401, USA
2 Institute of Cognitive Science, University of Colorado Boulder, Muenzinger D244 345 UCB, Boulder, CO 80309, USA

new network is trained with each training set. The objective of this step is to reach equal or better classifier performance with the new training sets generated with synthetic data.

6.1.1 Data Collection

In this research, a 3-D model of a hazardous roof condition is created using close-range digital photogrammetry. The location of the roof area and the roof beams mapped by the ground control personnel are shown in Figure 6.1. This location is selected based on the recommendation of the ground control personnel. These images are collected on August 3rd, 2020.

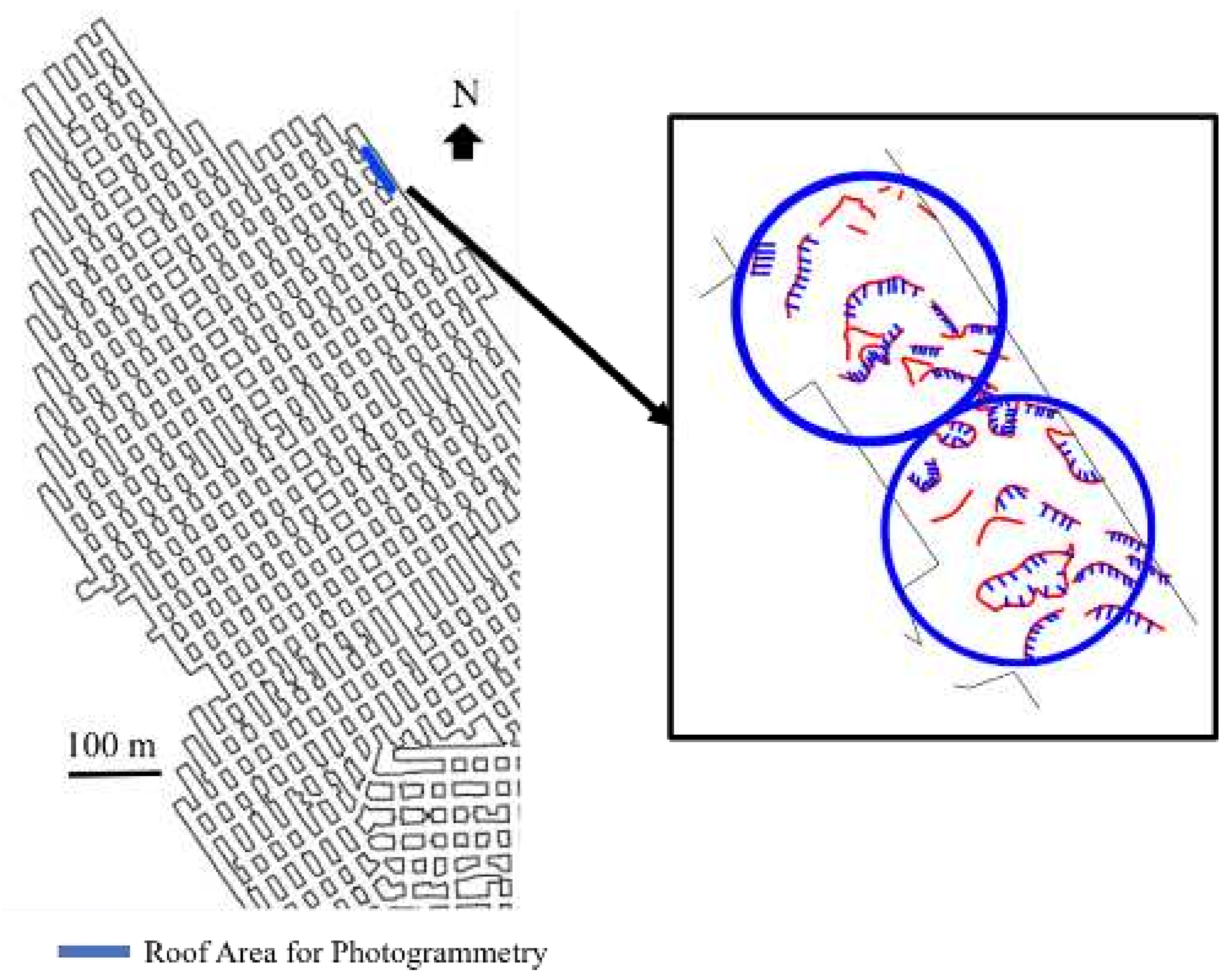


Figure 6.1 Location of hazardous roof area selected for photogrammetry

Roof beams for the hazardous roof area selected for photogrammetry are quantified in terms of depth and frequency using the techniques explained in Section 4.1. However, statistical analysis is not used since the area selected for photogrammetry is small for any meaningful statistical comparison. Instead, the summary statistics are compared with that of previous roof fall locations (Table 6.1).

Table 6.1 Summary statistics (Mean, Standard Deviation (SD)) of beam frequency and depth for previous roof fall locations and area chosen for photogrammetry

	Frequency		Depth	
	Previous roof fall	Hazardous Photogrammetry	Previous roof fall	Hazardous Photogrammetry
Mean	9.76	15	1.47	1.42
SD	3.76	2	0.65	0.14

The frequency of the roof beams in the area selected for photogrammetry is higher than the average roof beam frequency in previous roof fall locations, whereas the depth of the roof beams is slightly lower in the area selected for photogrammetry. Based on the summary statistics and the recommendation of the ground control personnel at the mine, the area chosen for photogrammetry is decided to be in the hazardous roof class. The number of photographs taken to be used in the photogrammetry is 270.

6.1.2 3-D Model Generation with Photogrammetry

Close-range digital photogrammetry uses overlapping images to create 3-D models with texture information. In photogrammetry, matching points between overlapping images are found. Based on this information, camera locations are estimated, and a sparse point cloud is built. Then, a single dense point cloud is generated by calculating the depth information for each camera. After the construction of the dense point cloud, it can be used to generate 3-D mesh models and texture files. In close-range photogrammetry, the relatively small distances between the camera and the surface allow the use of off-the-shelf DSLR cameras (Haneberg, 2008). Photogrammetry has been employed to create accurate 3-D models in many geotechnical engineering applications such as deformation detection (Zhang et al., 2017), joint mapping (Buyer et al., 2020), and slope stability (Bar et al., 2020).

The photogrammetry software used in this research is Agisoft Metashape. The generated 3-D model of a hazardous roof condition and its dimensions are shown in Figure 6.2.

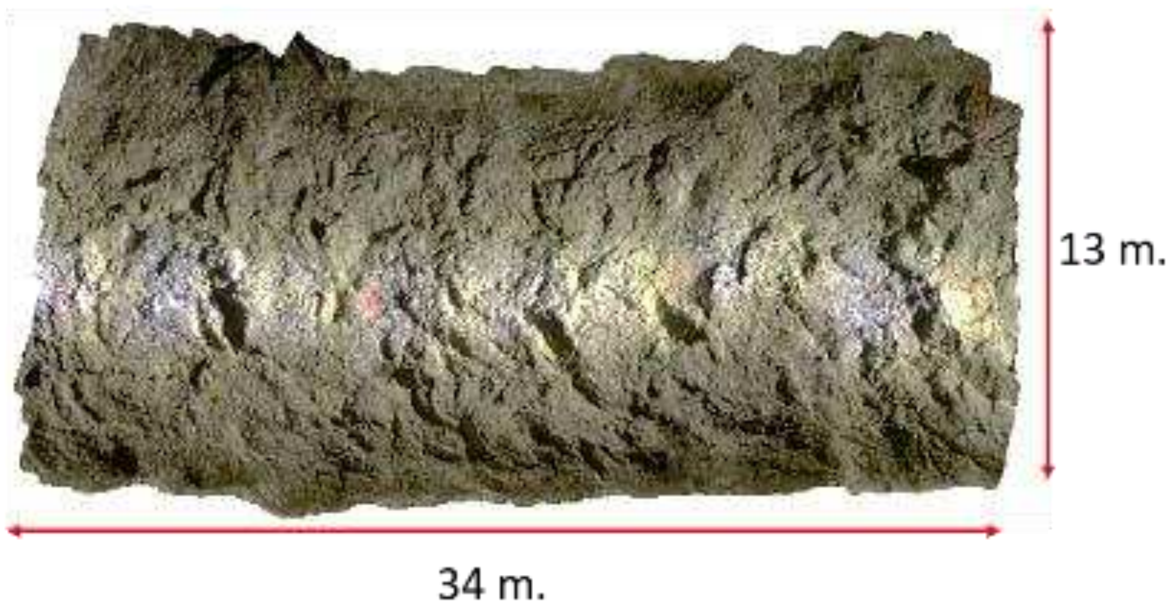


Figure 6.2 3-D Model generated with photogrammetry

6.1.3 Image Rendering

The generated 3-D model is transferred to Blender computer graphics software for image rendering. In the Blender scene, a camera that renders 224 x 224 pixels images as required by the CNN, a light source, and a camera rig for the camera to use a path are placed. The camera and the light source start moving together from the start point of the camera rig to the endpoint

(Figure 6.3). During this move, images of the hazardous roof condition are rendered on the 3-D model.

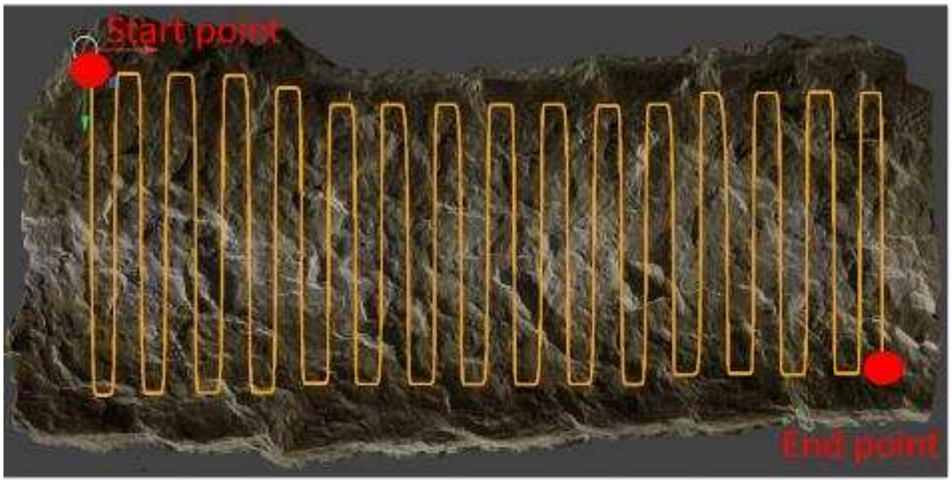


Figure 6.3 Blender scene

Using the photogrammetry and image rendering workflow described above, 360 synthetic hazardous roof images are rendered. Figure 6.4 shows an example of a synthetic hazardous roof image.

Figure 6.4 Example synthetic hazardous roof image

6.1.4 Dataset Generation with Synthetic Images

The synthetic images are used to replace, or augment real hazardous images collected on-site previously. Hazardous: non-hazardous image ratio is 1:2 in every data set created in this step to control for the effects of an unbalanced training set. Therefore, for each hazardous roof image in the training set, we included two real non-hazardous roof images. The validation set does not include any synthetic images. Table 6.2 shows the new training sets created in this step, and their distribution of real and synthetic images for each class.

Table 6.2 Datasets created with real and synthetic images

Dataset	Hazardous		Non-hazardous	
	Real	Synthetic	Real	Synthetic
Network (B)	0	266	532	0
Network (C)	133	266	798	0
Network (D)	266	133	798	0
Network (E)	266	266	1064	0
Validation	66	0	132	0

Training sets for the hazardous class are created with different ratios of real and synthetic images to examine the effects of synthetic images on network performance. CNNs are trained with each of these training sets using the same architecture and transfer learning approach described in Chapter 5. The network performances are compared using the true positive rate (sensitivity) (TPR) and false-positive rate (FPR). Ideally, an improved network performance would show increased TPR and decreased FPR values. Equations 6.1 and 6.2 show how these two parameters are calculated;

$$\mathrm{TPR} = \frac{\text{True Positive}}{\text{True Positive+False Negative}} \tag{6.1}$$

$$\mathrm{FPR} = \frac{\text{False Positive}}{\text{False Negative+True Negative}} \tag{6.2}$$

6.1.5 Network Performances Obtained After Synthetic Data Generation

A CNN network is created by training the fully connected layer with the new training sets created and their performances evaluated based on TPR and FPR. Figure 6.5 shows the TPR and FPR results of the CNN validation performances.

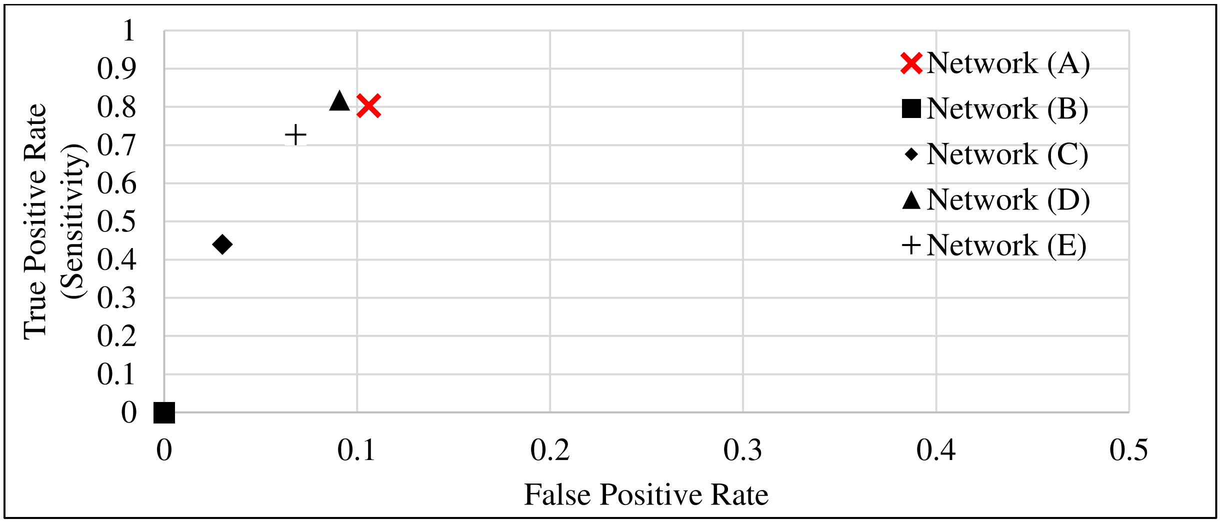


Figure 6.5 Network performances after synthetic data generation

In the above graph, the red X (Training set A) shows the performance of the network trained with only 798 real images. Here, the objective is to find an improvement on this network. Since TPR and FPR are employed as performance measures, improved network performance means increased TPR and decreased FPR. Therefore, the graph shows that the network trained with training set D shows better validation performance. The confusion matrix of the network trained with Training set D is given in Table 6.3.

Table 6.3 Confusion matrix of the network after synthetic data generation

		Predicted	
		Hazard	Non-hazard
Actual	Hazard	54	12
	Non-hazard	12	120

The confusion matrix shows that, after synthetic data generation, the classification accuracy of both classes has improved. In the next section, training set D is utilized to look for further performance improvements using data sampling, since it gives the best performance so far.

6.2 Data Sampling

In this section, a simple and easy-to-implement data sampling algorithm based on image features is introduced. The method rejects non-informative training samples for both classes. New training sets are created by applying the data sampling method at different levels, and a new network is trained with each of the new training sets. The objective of this chapter is to achieve better classifier performance after implementing the proposed data sampling method.

In this research, CNNs are trained to detect stress-induced roof beams which are indicators of a hazardous roof condition. The roof beams are repeated features in the same directions, and Gabor filters are effective in extracting linear features in selected orientations and wavelengths. A 2-D Gabor filter acts as a local band-pass filter with certain direction and wavelength characteristics. Typically, an image is filtered with a set of Gabor filters that have different directions and wavelengths. Gabor filters have been successfully applied to many image processing applications such as texture segmentation (Zhang et al., 2002), edge detection (Jiang et al., 2009), and feature extraction (Abhishree et al., 2015).

6.2.1 Details of the Data Sampling Technique

Gabor filters are used to identify different regions in hazardous and non-hazardous roof images. Therefore, first, a set of Gabor filters are designed. The orientations are between 0° and 150° in steps of 30°. The wavelengths are in increasing powers of two starting from 2 up to the hypotenuse length of the image. This set of orientations and wavelengths are based on the recommendations from Jain (1991) (Fig. 6.6).

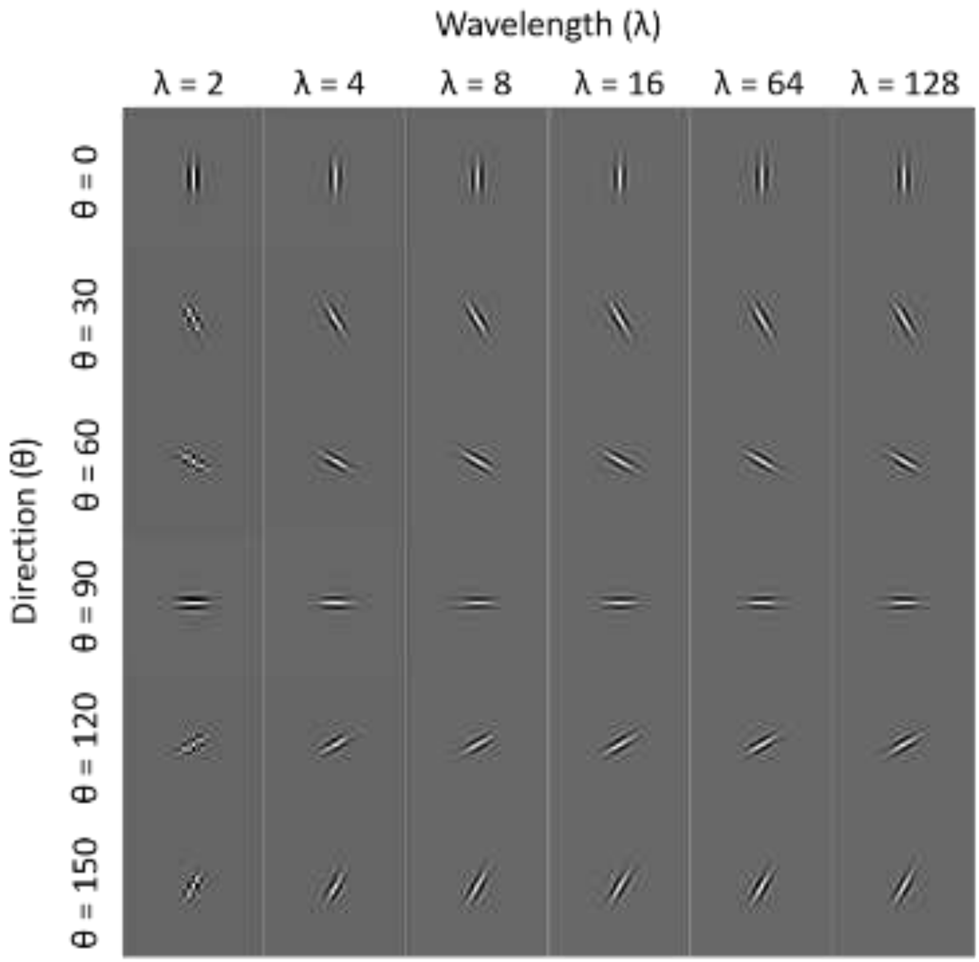


Figure 6.6 Gabor filters

Gabor magnitude response is extracted for each image using the filters shown above. To understand what Gabor magnitude response looks like in each image, principal component analysis is used to move from the 36-D representation of each pixel in an image into a 1-D magnitude value for each pixel. Figure 6.7 shows Gabor magnitude responses for a hazardous and non-hazardous roof image.[**]

[**] The source code is available on GitHub at https://github.com/erginisleyen/Isleyen-PhD-Research

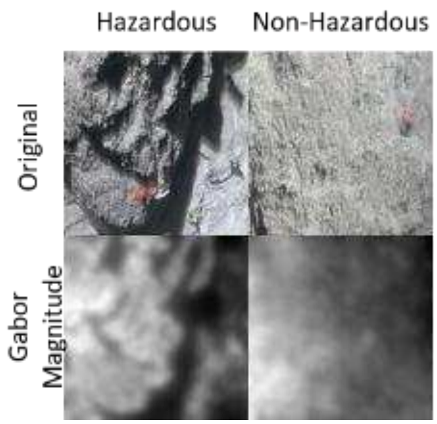


Figure 6.7 Examples of Gabor magnitude responses

The visualization of Gabor magnitude responses suggests that the hazardous roof class has more variance than the non-hazardous class in terms of Gabor magnitudes. Therefore, it is decided to further investigate the Gabor magnitude variance difference between the two classes. For this purpose, training dataset D is used since the network trained on this dataset gave the highest performance so far. The variance of the grayscale pixel values of each Gabor magnitude response image is calculated. A statistical t-test is used to test for a difference between the means of the Gabor magnitude variances. Summary statistics and p-values are given in Table 6.4.

Table 6.4 Summary statistics for Gabor magnitude variances (Mean, standard deviation (SD), degrees of freedom (df), t-values (t) and P-values (P-val)

	Hazardous	Non-hazardous
Mean	0.045	0.042
SD	0.01	0.01
df	1195	
t	4.37	
P-val	<.001	

The p-value between hazardous and non-hazardous roof images is smaller than the level of significance (0.05). The t-test concludes that the difference in Gabor magnitude variances that was found, or an even larger difference, would occur if the variance values were not different (null hypothesis) is less than the level of significance. Also, the summary statistics show that the hazardous roof images usually have higher Gabor magnitude variance values, whereas the non-hazardous roof images usually have lower Gabor magnitude variance values. It is therefore plausible to design a data sampling technique based on Gabor magnitude variances.

The idea behind the data sampling technique is classifying certain groups of images as outliers and removing them from the input dataset. For the hazardous class, hazardous roof images that have low Gabor magnitude variances are classified as outliers since hazardous roof images usually have higher Gabor magnitude variances compared to the non-hazardous class. Similarly, for the non-hazardous class, non-hazardous roof images that have high Gabor magnitude variances are classified as outliers since non-hazardous roof images usually have lower Gabor magnitude variances compared to the hazardous class. To categorize the images, we used the Gabor magnitude variance histograms of both classes (Figures 6.8 & 6.9).

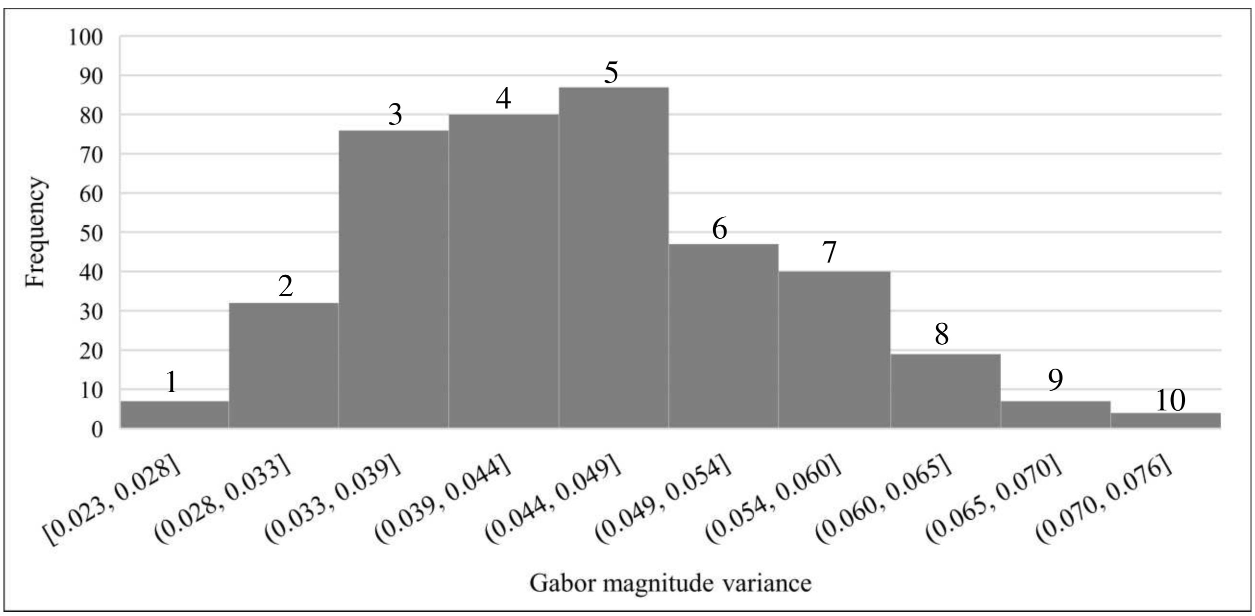


Figure 6.8 Histogram of Gabor magnitude variances for hazardous class

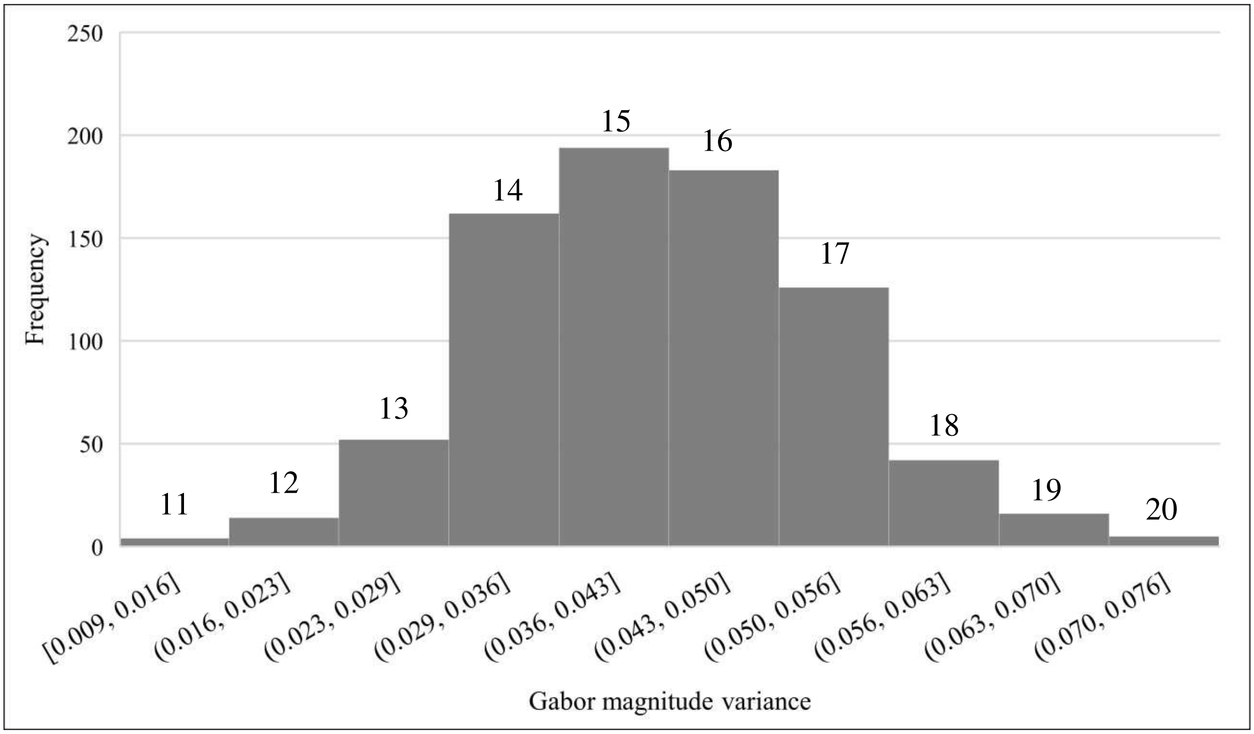


Figure 6.9 Histogram of Gabor magnitude variances for non-hazardous class

The histograms show that hazardous and non-hazardous images are both categorized into 10 groups. Each category is assigned a name that is shown at the top of the bars (e.g. 1,2,3…). In this data sampling technique, outlier categories are removed one by one from the input datasets to improve the quality of the training data. For the hazardous class, categories that have low Gabor magnitude variances are removed. For the non-hazardous class, categories that have high Gabor magnitude variances are removed. For example, images in category 20 have high Gabor magnitude variances even though they belong to the non-hazardous class. It is expected that removing these images from the dataset will improve the quality of training data, which may improve the network classification performance. Table 6.5 shows which categories are included in each training dataset created using the data sampling technique.

Table 6.5 Descriptions of training sets of each network after data sampling

Dataset	Hazardous Categories	Non-hazardous Categories
Network (D)*	1,2,3,4,5,6,7,8,9,10	11,12,13,14,15,16,17,18,19,20
Network (F)	2,3,4,5,6,7,8,9,10	11,12,13,14,15,16,17,18,19
Network (G)	3,4,5,6,7,8,9,10	11,12,13,14,15,16,17,18
Network (H)	4,5,6,7,8,9,10	11,12,13,14,15,16,17
Network (I)	2,3,4,5,6,7,8,9,10	11,12,13,14,15,16,17
Network (J)	2,3,4,5,6,7,8,9,10	11,12,13,14,15,16,17,18
Network (K)	3,4,5,6,7,8,9,10	11,12,13,14,15,16,17,18,19
Network (L)	4,5,6,7,8,9,10	11,12,13,14,15,16,17,18,19

*Training set D was used as the basis of the data sampling technique. The images in the training set D are not sampled.

The data sampling technique was applied to training set D which includes both real and synthetic images. All the other training sets listed in Table 6.5 are created by sampling training set D at different levels. Figure 6.10 visualizes how these training sets are created using the Network (G) as an example.

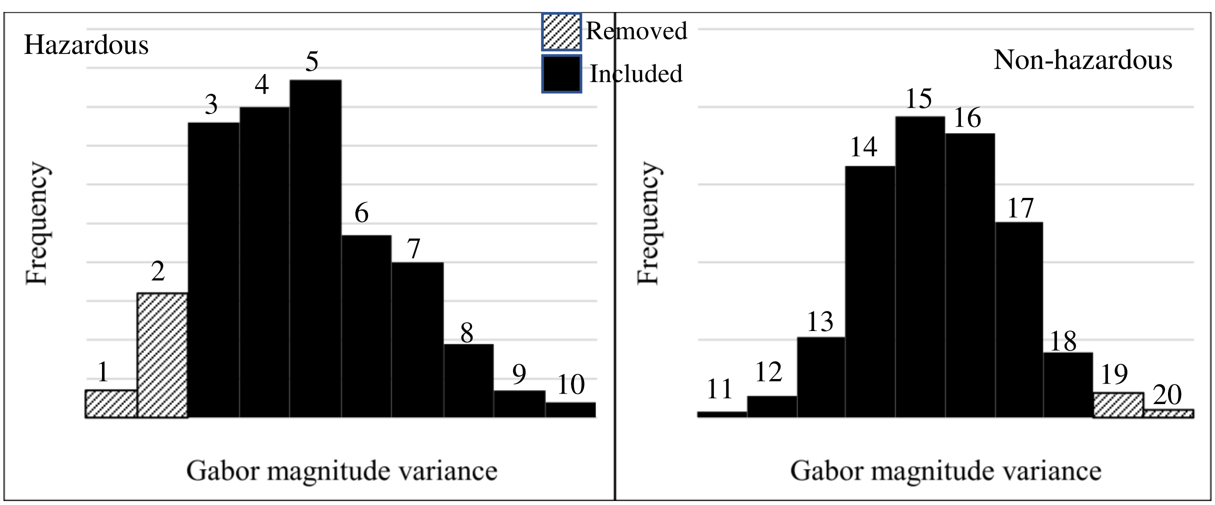


Figure 6.10 Example implementation of data sampling technique (Network G)

To create the training set of Network G, images in categories one and two are removed from the hazardous training set, and images in categories 19 and 20 are removed from the non-hazardous training set. The training sets of the other networks are prepared with the same logic. The distribution of real and synthetic images in each of the newly created training sets is given in Table 6.6.

Table 6.6 Distribution of real and synthetic images used for training each network after data sampling

Dataset	Hazardous		Non-hazardous	
	Real	Synthetic	Real	Synthetic
Network (F)	259	133	793	0
Network (G)	239	124	781	0
Network (H)	182	105	729	0
Network (I)	259	133	729	0
Network (J)	259	133	781	0
Network (K)	239	124	793	0
Network (L)	182	105	793	0

A CNN is trained using each of these training sets using the same architecture and transfer learning approach described in Chapter Five.

6.2.2 Network Performances Obtained with Data Sampling

The network performances are compared using the true positive rate (sensitivity) (TPR) and false-positive rate (FPR). Figure 6.11 shows the TPR and FPR results of the CNN validation performances.

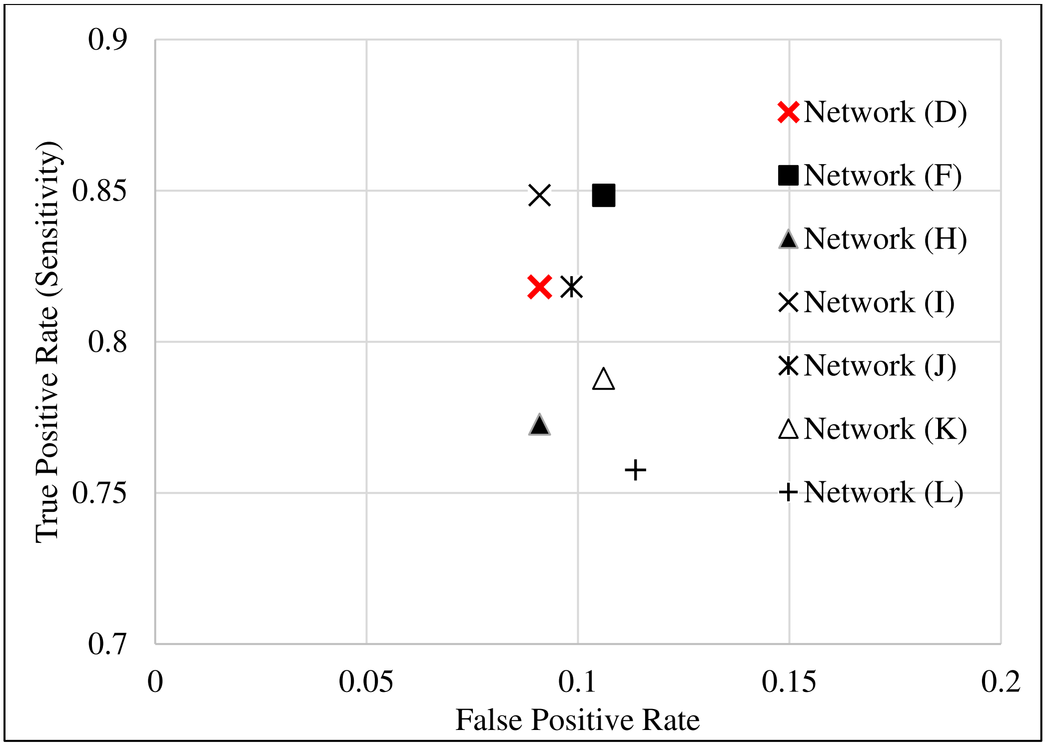


Figure 6.11 Network performances after data sampling

In the above graph, the red X (Training set D) shows the performance of the network trained in the previous step after synthetic data generation. Here, the objective is to find an improvement on this network. Since TPR and FPR are employed as performance measures, improved network performance means increased TPR and decreased FPR. Therefore, the graph shows that the network trained with the training set I shows better validation performance. Table 6.7 shows the confusion matrix of the network trained with training set I.

Table 6.7 Confusion matrix of the network after data sampling

		Predicted	
		Hazard	Non-hazard
Actual	Hazard	56	10
	Non-hazard	12	120

The confusion matrix shows that, after the data sampling step, the classification accuracy of the hazardous class improved while the classification accuracy of the non-hazardous class remained the same.

6.3 Results and Discussion

In this chapter, two methods are used to improve the classification performance of the AI-based roof fall hazard detection system, namely synthetic data generation and data sampling. First, the classifier results obtained after synthetic data generation step is investigate. Figure 6.12 shows the distribution of real and synthetic images in both classes in each training set created after synthetic data generation. The validation performance obtained with the classifiers trained with each of these training sets is also shown in Figure 6.12 (page 59), in terms of sensitivity and specificity. Sensitivity measures a classifier's ability to predict positive cases (i.e. hazardous) within the dataset, and specificity measures a classifier's ability to predict negative cases (i.e. non-hazardous).

Network A refers to the classifier trained with only real images in Chapter 5. The objective is to find an improvement on this network using synthetic images. Replacing the real hazardous images with synthetic hazardous images caused the classifier to predict everything as non-hazardous (Network B), thereby causing 100% specificity and 0% sensitivity. When the majority of images are synthetic in hazardous class (Network C), predictions are still mostly in favor of the non-hazardous class. However, sensitivity increases significantly when the hazardous training set is supplemented with synthetic data (Network D). In this case, both sensitivity and specificity are higher than those obtained by a classifier trained with only real images (Network A). Further increasing the number of synthetic images in Network E causes sensitivity to decrease. The increase in specificity in Network E is likely to be caused by the increased number of images in the non-hazardous class. This analysis shows that classifier performance is improved by adding synthetic images in hazardous class. The best performance was achieved when the real images are still the majority. Also, keeping the ratio of the number of images between the classes the same by adding more non-hazardous images was effective in maintaining high specificity.

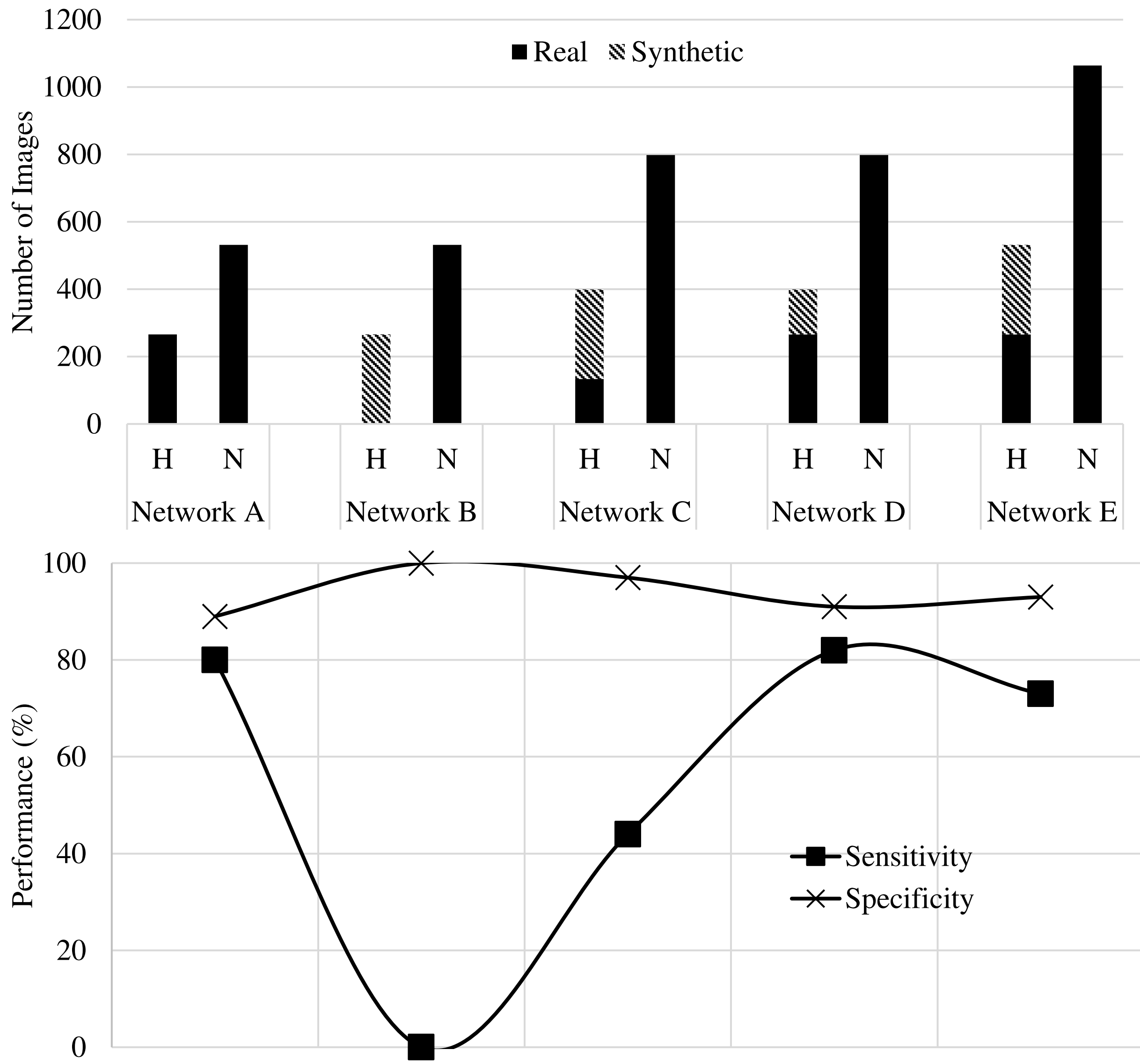


Figure 6.12 Training sets after synthetic data generation and performance metrics (H: Hazardous, N: Non-hazardous)

The results show that replacing real hazardous data with synthetic data causes the network to completely fail in predicting hazardous roof conditions. However, augmenting the real data with synthetic data improves the network performance for both hazardous and non-hazardous class. Therefore, in cases where real data collection for hazardous conditions is too dangerous for the mine personnel, the data that represent the pattern of the current hazard status can be generated synthetically using techniques that prevents the personnel from exposing themselves to hazards. However, it should be noted that the photogrammetry is only used to describe the synthetic data generation technique as a proof-of-concept. Since photogrammetry also requires on-site data collection, it is not an ideal technique in terms of safety. Sensors such as lidar, allow generating 3-D models much faster or with remote control. Implementing these types of sensors with the described methodology could generate more synthetic images from different areas of the mine in a shorter time. Therefore, it is possible that increased number of synthetic images with more variety could further improve these results.

Second, the classifier results obtained after the data sampling step is investigated. Figure 6.13 (page 61) shows the distribution of images in the training sets created in this step, and the specificity and the sensitivity of the networks trained with these training sets.

Network D refers to the classifier with the best validation performance after the synthetic data generation step. The objective is to find a network with better performance after implementing the data sampling technique. In hazardous class, the images with low Gabor magnitude variance are removed gradually, whereas, for non-hazardous class, the images with high Gabor magnitude variance are removed gradually. The specificity shows a slight increase when the non-hazardous images with high Gabor magnitude variances are removed from the training set. However, sensitivity only increases after removing five images with the lowest Gabor magnitude variances. Removing more images from the hazardous training set, causes sensitivity to decrease. Based on these observations, the Network I show the best performance. It has higher specificity and sensitivity than those obtained from the network after the synthetic data generation step. Figure 6.14 (page 62) shows the training and validation accuracy of the Network I.

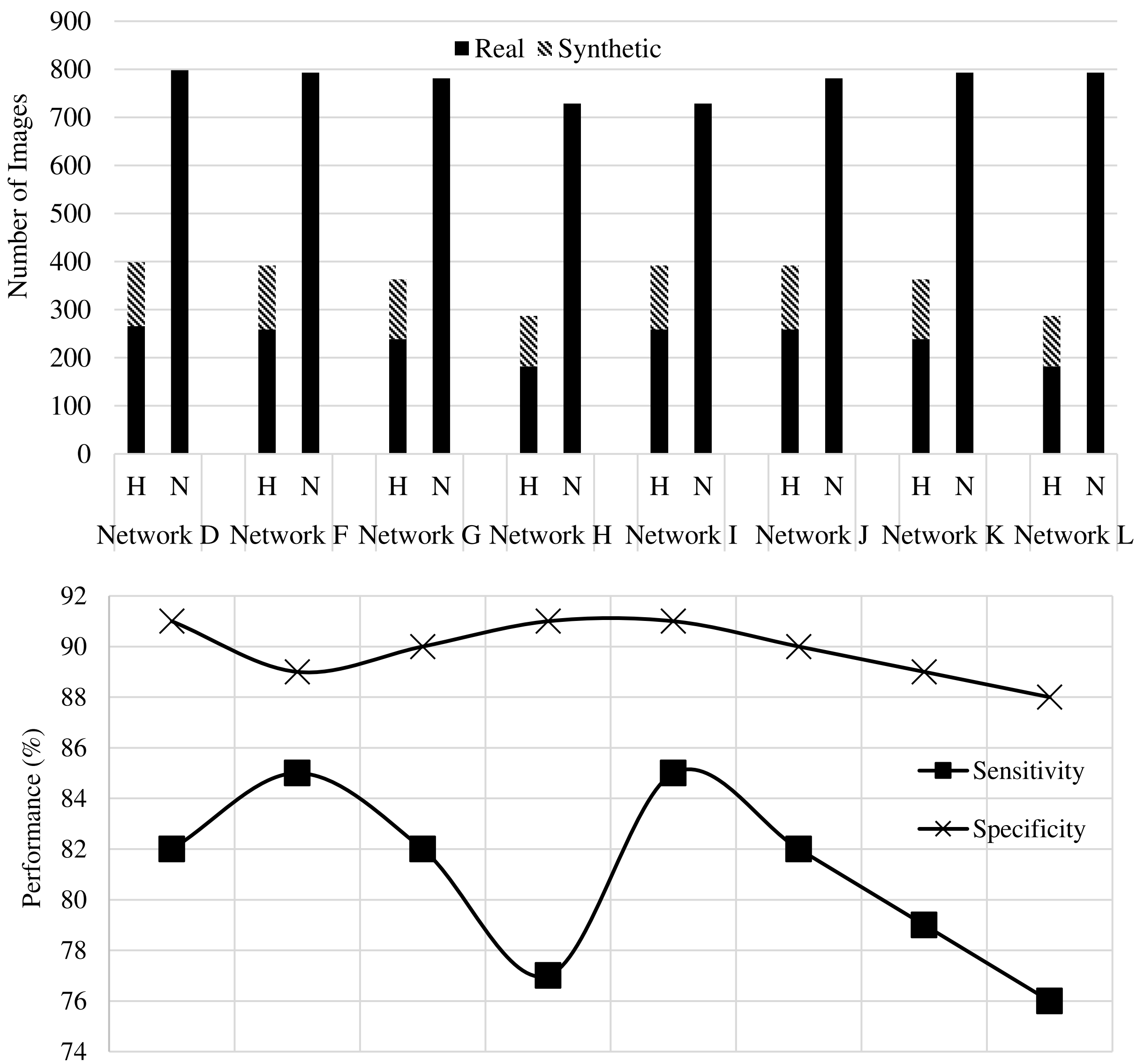


Figure 6.13 Training sets after data sampling and performance metrics (H: Hazardous, N: Non-hazardous)

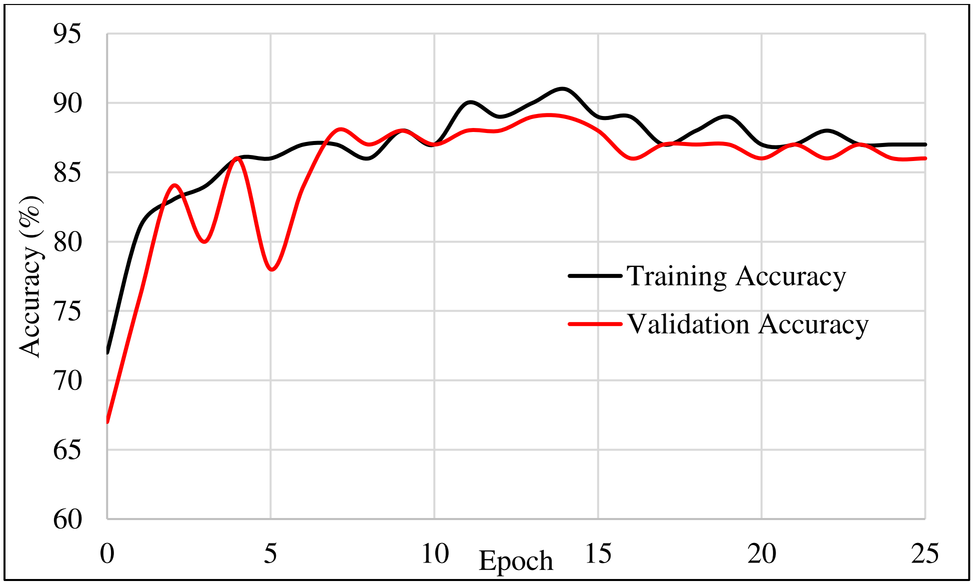


Figure 6.14 Training and validation accuracy

The validation dataset is used to obtain the validation accuracy of the network. In addition to validation accuracy, the training accuracy is calculated: this is the accuracy of the model when applied to the training images. Training accuracy is used as an indicator of overfitting which means that the model fits too closely with the training data and fails to generalize for the prediction of future data sets. In Figure 6.14, the general trend of training accuracies is similar to the validation accuracies. This shows that the model does not overfit to the training data.

CHAPTER 7

TESTING AND INTERPRETATION OF THE CNN

Submitted to the Geoscience Frontiers Journal

Ergin Isleyen[††1], Sebnem H. Duzgun[1], McKell R. Carter[2]

In this chapter, first, a comparison of the classifier performance after each step is given. Then, the classifier performance is validated using a test dataset. Lastly, a deep learning interpretation technique is applied to understand the similarities and differences between the visual features used by the expert and the CNN to make a classification decision.

Evaluation of performance metrics is important to understand how well the network is performing. Different performance metrics have been used to measure classifier's ability for different conditions. Sensitivity is the ability of the network to correctly identify hazardous conditions. Specificity is the ability of the network to correctly identify non-hazardous conditions. Precision is the probability that, given a hazardous condition, the sample is hazardous. F1 score is the harmonic mean between sensitivity and precision. Therefore, it is a better performance measure than accuracy in cases where there is class imbalance. The performance of the final network developed in this thesis is investigated using these metrics. Table 7.1 compares the network performances after each step, in terms of sensitivity, and specificity, precision and F1 score.

[††] Corresponding author. Direct correspondence to isleyen@mines.edu
1 Mining Engineering Department, Colorado School of Mines, 1500 Illinois St., Golden, CO 80401, USA
2 Institute of Cognitive Science, University of Colorado Boulder, Muenzinger D244 345 UCB, Boulder, CO 80309, USA

Table 7.1 Comparison of classifier performances after each step

Step	Sensitivity (%)	Specificity (%)	Precision (%)	F1 Score
Transfer Learning	80	89	79	79.5
Synthetic Data Generation	82	91	82	82
Data Sampling	85	91	82	83

The table shows that the classification accuracy, sensitivity, and F1 score increases after each step, whereas specificity and precision increases after synthetic data generation and remains the same after data sampling. This verifies the success of the proposed methodology as it is tested on the validation set. However, it is also important to verify the classifier performance on an independent test dataset.

7.1 Data Collection

The network with the best performance is tested on an independent sample to verify the performance of the classifier. The test images are collected from areas reserved for test image data collection and no training, validation, or photogrammetry image was collected in those areas. Figure 7.1 (page 65) shows the locations where test images are taken. The test images are collected on August 4th, 2020.

Roof beams for the hazardous and non-hazardous roof areas selected for testing are quantified in terms of depth and frequency using the techniques explained in Section 4.1. The frequency and depth of roof beams between hazardous and non-hazardous testing locations are compared statistically. A statistical t-test is used to test for a difference between the means of the two groups for each feature. Summary statistics and p-values are given in Table 7.2 (page 66).

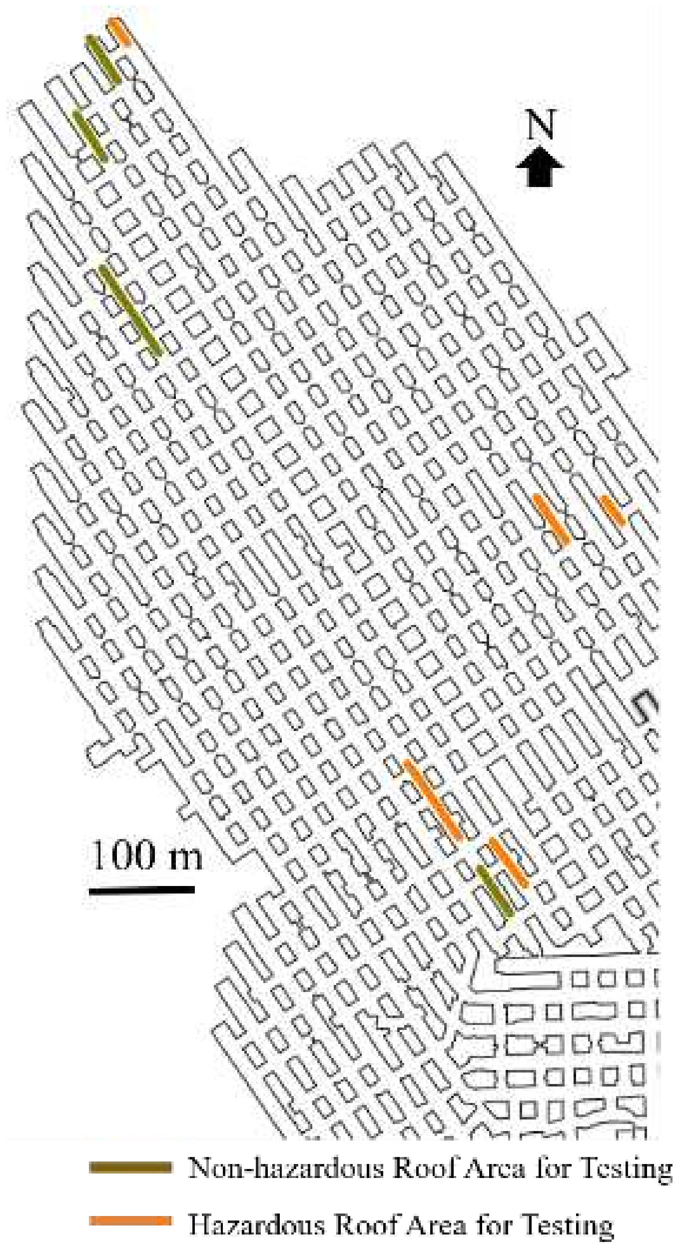


Figure 7.1 Data collection locations for testing

Table 7.2 Summary statistics and p-values for test dataset

	Frequency		Depth	
	Hazardous	Non-hazardous	Hazardous	Non-hazardous
Mean	13	4.4	1.7	1.0
SD	1.9	2.8	0.23	0.81
df	21		21	
t	8.4		2.6	
P-val	<.001		.02	

For both frequency and depth values of roof beams, the P-values between "hazardous test" and "non-hazardous test" locations are smaller than the level of significance (0.05). The t-test concludes that the difference that was found in frequency and depth between roof beams around hazardous test locations and non-hazardous test locations, or an even larger difference, would occur if the roof conditions were the same (null hypothesis) is less than the level of significance. Therefore, it is plausible that these locations can be used as data collection locations for testing the final network performance. Compared to frequency and depth values of training data collection locations (Table 5.1), frequency and depth values for both hazardous and non-hazardous test locations are slightly higher.

7.2 Testing the Network

In order to verify the performance of the final network which was obtained after the data sampling stage (Network I), it is tested on an independent test dataset. The test dataset includes 80 hazardous and 160 non-hazardous roof images. The confusion matrix of the test results is given in Table 7.3.

Table 7.3 Confusion matrix for testing

		Predicted	
		Hazard	Non-hazard
Actual	Hazard	79	1
	Non-hazard	24	136

The overall test accuracy is 90%, and the confusion matrix shows that the network is very successful in classifying hazardous roof images. The test accuracy on hazardous class is even higher than the validation accuracy on hazardous class. The performance boost can be explained by the higher number of frequency and depth of roof beams around the hazardous test locations, making it easier for the network to identify features that indicate hazardous roof conditions. However, the classification accuracy shows a slight decrease when it comes to predicting non-hazardous roof images. This can also be explained by the higher number of frequency and depth of roof beams around the non-hazardous test locations, making it more difficult for the network to identify the smooth regions that indicate non-hazardous roof conditions. In terms of roof fall hazard detection, the cost of a false positive is much less than the cost of a false negative. Therefore, it is concluded that the network performance on test dataset is acceptable.

7.3 Interpretation of the Network Predictions

Deep learning models are typically “black boxes”, meaning that the user does not receive a reasoning in terms of input features for the predictions made by the model, regardless of how accurate the predictions are. However, if there are actions to be taken based on the predictions of the network, as in the case of roof fall hazard detection, understanding the reasons behind the predictions carries great importance for the user. Understanding the features used by the networks increases the confidence of the user, which is critical for managing hazards that could cause severe injuries and fatalities.

Since CNN takes images as input, the interpretability of CNN models can be achieved by producing visual explanations. Several techniques have been developed to interpret the predictions made by CNN. These techniques usually create visualizations that highlight image regions that contribute most to network predictions at the pixel level (Sundararajan et al., 2017; Selvaraju et al., 2017; Shrikumar et al., 2017) or object-level (Zhang et al., 2018).

7.3.1 Details of the Deep Learning Interpretation Technique

In this study, an attribution technique called "integrated gradients" is used for attributing the predictions of CNN to the input images (Sundararajan et al., 2017). In this technique, a series of images are interpolated, increasing in intensity, between a black baseline image and the original image. The network predicts the class of each interpolated image and calculates the confidence scores of each prediction. The gradients of the features are calculated with respect to the model output. Integral approximation is used to accumulate the gradients. The integrated gradients technique is pixel-based; therefore, the outputs show the most important pixels in each image for the hazard or non-hazard prediction. Some examples of the integrated gradients technique are given in Figure 7.2.

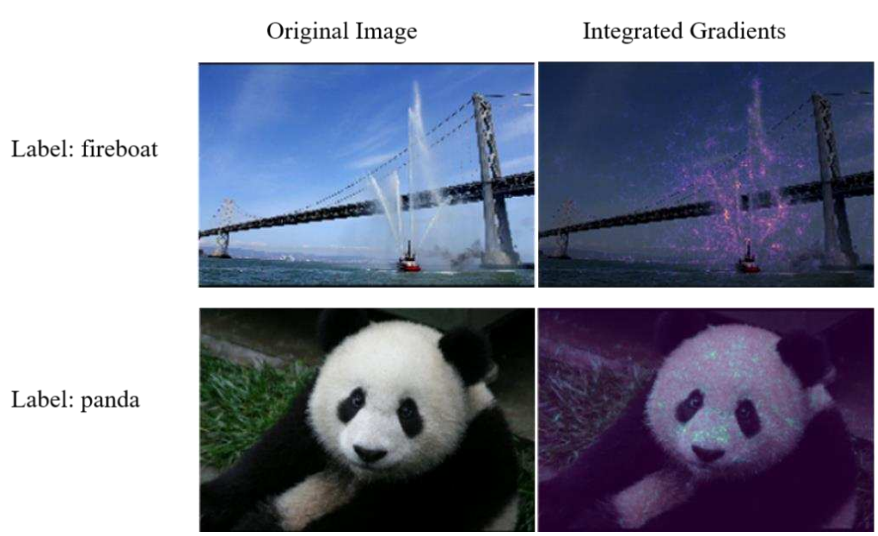


Figure 7.2 Examples of integrated gradients results (Source: TensorFlow, n.d., CC by 4.0)

The integrated gradients method is used to study pixel importance in predictions made by a network. The gradients are computed for the output of the highest-scoring class with respect to the pixels of the input image. The example images show that integrated gradients are very effective at reflecting the distinctive features of an input image.‡‡

7.3.2 Results of the Integrated Gradients Technique

To understand the visual features used by the system for classification decisions, an integrated gradient technique is used on every image in the test dataset. Figure 7.3 (page 70) shows examples of the results of integrated gradients.

For hazardous roof conditions, integrated gradients show that the network's prediction is based on texture changes from light colors to darker colors. Texture changes are the results of changes in depth of roof formations where high horizontal stresses have created roof beams. For non-hazardous roof conditions, integrated gradients show that the network's prediction is based on regions where the texture colors remain stable. These regions exhibit smooth roof conditions. The results of integrated gradients on test images are given in Appendix A.

7.4 Results and Discussion

In this chapter, the performance of the classifier is verified with the test dataset. It was shown previously that the accuracy of the classifier on the test dataset is 90%. Here, the performance of the classifier is evaluated with other metrics that are given in Table 7.4.

Table 7.4 Performance metrics obtained from the test dataset.

Sensitivity	99%
Specificity	85%
Precision	77%

‡‡ The source code is available on GitHub at https://github.com/erginisleyen/Isleyen-PhD-Research

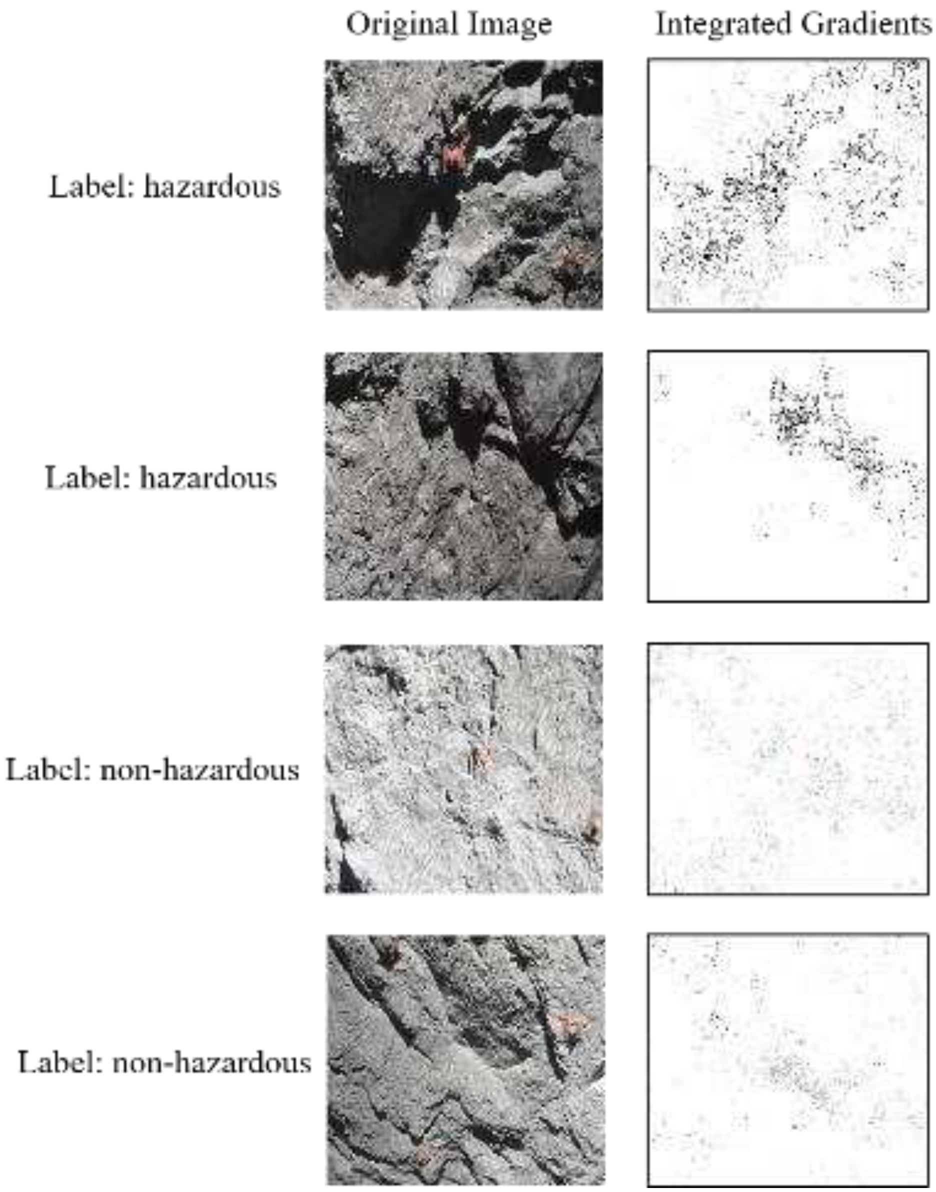


Figure 7.3 Examples of the results of the integrated gradients

Sensitivity calculates the percentage of actual hazardous images the classifier correctly identified, and the specificity calculates the percentage of actual non-hazardous images the classifier correctly identified. Since the consequences of a false negative could be dangerous in hazard detection, sensitivity is a more important measure than specificity. The sensitivity value obtained from testing is very high. Therefore, the classifier developed with the methodology proposed in this study is a promising hazard detection tool. Also, precision evaluates how often the classifier was correct in its hazardous predictions. The number of false positives decreases the precision of the model. In hazard detection, the cost of a false positive is much less than a false negative. Therefore, the precision value obtained from testing is acceptable.

Investigating the results of the integrated gradients allows the user to understand the visual features used by the network to make a classification decision. The CNN network for roof-fall-hazard detection uses pixels around the roof beams for hazardous roof condition predictions, whereas for the non-hazardous-roof-condition predictions it uses the pixels around the smooth areas. The features used by CNN is similar to the experts' description of their intuitive decision-making skills in identifying features of hazardous and non-hazardous roof conditions. In this respect, the system mimics the expert's judgment on hazard detection. This work provides an important example of providing human-understandable justification of network predictions to create a transparent hazard management tool.

CHAPTER 8

CONCLUSIONS AND RECOMMENDATIONS FOR FUTURE RESEARCH

AI-based autonomous technologies have the potential to transform the mining industry by improving operational efficiency and safety performance, boosting productivity, and cutting costs. However, the transition to fully autonomous mining requires various changes throughout the workplace including how geotechnical hazards are managed. Developing deep learning systems require large amounts of data. In the case of autonomous geotechnical hazard management, this requirement becomes a major obstacle since the data needs to be collected on-site for geotechnical hazard applications. On-site data collection limits the deep learning applications in geotechnical hazard management. Some of the reasons behind this limitation are site access restrictions due to safety concerns and operational requirements, additional costs associated with equipment required to collect large amounts of data, and long work hours needed for on-site data collection. This thesis presents a methodology to develop an AI-based autonomous geotechnical hazard management system with small datasets. The methodology includes transfer learning, synthetic data generation, and data sampling techniques. The proposed methodology is implemented on a case study to develop a CNN classifier for a roof fall hazard detection using a small dataset collected on-site.

Roof falls caused by high horizontal stresses are critical problems in limestone mines in the Eastern and Midwestern United States. Apart from fundamental hazard management techniques such as roof bolting and stress control layout, roof fall hazards are managed by daily visual inspections. In the Subtropolis limestone mine, these visual inspections are carried out by ground control personnel who look for horizontal stress-induced roof beams. In this study, a

CNN-based, autonomous roof-fall-hazard detection system was developed for the Subtropolis case.

The system was trained with images showing hazardous and non-hazardous roof conditions in Subtropolis mine, followed by a transfer learning approach. Using transfer learning, only the fully connected layer of the CNN was trained with images showing hazardous and non-hazardous conditions. The network was pre-trained with ImageNet. The accuracy obtained by the network (86%) is higher than the base accuracy which is 66% (assuming every image is labeled as non-hazardous). The results confirm that low-level features provided by pre-training provide a suitable starting point for training the fully connected layer even in a distant task. This research demonstrated that transfer learning is a useful approach to develop an AI-based roof fall hazard detection system with small datasets.

The classifier accuracy was improved by creating new training sets with augmenting real images with synthetic images. New networks are trained with the new training sets. After augmenting real images with synthetic images, classification accuracy for both hazardous and non-hazardous class increased compared to the network trained only with transfer learning. The results showed that replacing real images with synthetic images causes the classifier to fail but augmenting real images with synthetic images produced better classification accuracy. This approach is useful when collecting more images for hazardous roof conditions is not possible due to safety reasons. In such cases, real data collection step can be replaced with a synthetic data generation step. However, it should be noted that the photogrammetry may not be practical since it also requires on-site data collection. In this thesis, photogrammetry was used to test the feasibility of using synthetic images. Use of sensors that can generate 3-D models much faster and have the option to be controlled remotely is more suitable.

An easy-to-implement data sampling technique is introduced that is based on the differences between Gabor magnitude responses of images. New training sets are created by applying the data sampling technique at different levels. The technique improves the dataset quality by removing non-informative images from the dataset. The implementation of the data sampling technique further improved classifier performance. After implementing the data sampling technique, the classification accuracy of hazardous class increased while the accuracy of non-hazardous class remained the same compared to the network trained with a training set created by augmenting real images with synthetic images and a transfer learning approach.

Classification results are interpreted using a deep learning interpretation technique called integrated gradients. Majority of the roof fall hazard management techniques proposed in the literature utilize expert's engineering judgment. This shows that expert input is valuable, and an autonomous hazard detection system should be able to use the same visual features as the expert to build user confidence. In this thesis, interpretation of the classifier predictions shows that the system utilizes the same visual features to make a classification decision on roof fall hazard detection.

The mining and tunneling industries expect the increasing use of autonomous systems. Detecting geological hazards without human involvement is an important milestone to achieve fully autonomous mining operations. This study presents the first step toward autonomous geological hazard detection in the mining and underground construction industry. Combining the network developed in this study with robotic systems or autonomous vehicles can provide continuous and real-time hazard detection.

In cases where hazard management depends on visual inspections, ground control personnel are exposed to roof fall hazards daily. Autonomous hazard detection tools eliminate these risks since human input is only required during initial system development. The role of the expert, then, becomes that of making the decisions and determining actions to be taken based on the predictions made by the system. Experts are no longer exposed to roof fall safety hazards.

Intuitive skills for hazard detection are learned by experience. It means that training new personnel for roof hazard detection takes extensive effort and time. Also, site-specific knowledge may be lost when an experienced employee leaves. This brings serious safety risks to the personnel still working in the mine. The AI-based system proposed in this study achieves high-accuracy in replicating expert judgment on hazard detection and can perform at a stable accuracy within the same geological and operational conditions. Therefore, AI-based hazard detection systems could help prevent the loss of expert knowledge that has been created through years of experience and prevents safety risk in the case of expert unavailability.

The limitation of this research is that the roof fall hazard detection system uses the ground control personnel' labeling of hazardous and non-hazardous conditions as ground truth. Therefore, it can be argued that the system classification accuracy (90%) can only be as good as an expert classification. However, the quantitative analysis of the relationship between the expert's labeling and roof beams showed that the expert classification is highly accurate.

Recommendations for future work are listed below;

- The system presented in this thesis was developed to address the roof fall problems of a single mine. Expanding the training set with data from multiple sites can increase the generalization ability of the system and improve the results.
- The CNN architecture used in this thesis was pre-trained on the ImageNet dataset. Although the previous uses of ImageNet in the literature and this research showed that it is highly successful at providing the low-level features for distant classification tasks,

using a network pre-trained on a geotechnical classification task may improve the classifier performance.

- Recent research indicates that transfer-learning models that initialize training from different pre-training checkpoints may outperform models whose transfer training is limited to the final layer (Neyshabur et al., 2020).
- For synthetic data generation, using a remote sensing system such as lidar instead of photogrammetry can assist in generating 3-D models faster. In this way, larger and more diverse synthetic datasets can be created.
- For synthetic data generation, a variational auto-encoder algorithm can be used. A variational autoencoder can be used to generate new images that are similar to the training set but not the same.
- To improve the success of the data sampling technique, other image features can be utilized. Gabor magnitudes were shown to be useful as it is based on linear features in the image. A variation of the data sampling technique can be created by using a combination of Gabor magnitudes with other textural information.

REFERENCES

Abhishree, T. M., Latha, J., Manikantan, K., & Ramachandran, S. (2015). Face recognition using Gabor filter based feature extraction with anisotropic diffusion as a pre-processing technique. *Procedia Computer Science, 45*, 312-321.

Ando, S., & Huang, C. Y. (2017). Deep over-sampling framework for classifying imbalanced data. In M. Ceci, J. Hollmen, L. Todorovski, C. Vens, & S. Dzeroski (Ed.), *Machine Learning and Knowledge Discovery in Databases. ECML PKDD 2017. Lecture Notes in Computer Science. 10534.* Springer.

Bar, N., Kostadinovski, M., Tucker, M., Byng, G., Rachmatullah, R., Maldonado, A., Potsch, M., Gaich, A., McQuillan, A., Yacoub, T. (2020). Rapid and robust slope failure appraisal using aerial photogrammetry and 3D slope stability models. *International Journal of Mining Science and Technology, 30*(5), 651-658.

Bertoncini, C. A., & Hinders, M. K. (2010). Fuzzy classification of roof fall predictors in microseismic monitoring. *Measurement, 43*, 1690-1701.

Buyer, A., Aichinger, S., & Schubert, W. (2020). Applying photogrammetry and semi-automated joint mapping for rock mass characterization. *Engineering Geology, 264*.

Chen, J., Yang, T., Zhang, D., Huang, H., & Tian, Y. (2021). Deep learning based classification of rock structures of tunnel face. *Geoscience Frontiers, 12*(1), 395-404.

Clareo and the World Innovation Network (TWIN). (2019). *Accessing the Fast and Furios Pace of Autonomy to Transform Mining.* Retrieved from https://clareo.com/wp-content/uploads/2020/05/Clareo-TWIN-Autonomy-Whitepaper-0619.pdf

Cunha, A., Pochet, A., Lopes, H., & Gattass, M. (2020). Seismic fault detection in real data using transfer learning from a convolutional neural network pre-trained with synthetic seismic data. *Computers and Geosciences, 135*.

Darling, P. (2011). *SME mining engineering handbook.* Englewood, CO: Society for Mining, Metallurgy, and Exploration.

Deng, J., Dong, W., Socher, R., Li, L. J., Li, K., & Fei-Fei, L. (2009). ImageNet: a large-scale hierarchical image database. *IEEE Conference on Computer Vision and Pattern Recognition*, (pp. 248-255). Miami, FL.

Department of Mines and Petroleum. (2015). *Safe mobile autonomous mining in Western Australia - code of practice.* Resources Safety, Department of Mines and Petroleum, Western Australia.

Dolinar, D. R. (2003). Variation of horizontal stresses and strains in mines in bedded deposits in the eastern and midwestern United States. In S. S. Peng, C. Mark, A. W. Khair, & K. A. Heasley (Ed.), *Proceedings of the 22nd International Conference on Ground Control in Mining* (pp. 178-185). Morgantown, WV: West Virginia University.

Douzas, G., & Bacao, F. (2017). Self-organizing map oversampling (somo) for imbalanced data set learning. *Expert Systems with Applications, 82*, 40-52.

Duzgun, H. (2005). Analysis of roof fall hazards and risk assessment for Zonguldak coal basin underground mines. *International Journal of Coal Geology, 64*(1-2), 104-115.

Duzgun, H. S., & Einstein, H. H. (2004). Assessment and management of roof fall risks in underground coal mines. *Safety Science, 42*(1), 23-41.

Esterhuizen, G. S., Dolinar, D. R., & Iannacchione, A. T. (2008). Field observations and numerical studies of horizontal stress effects on roof stability in US limestone mines. *Journal of the South African Institute of Mining and Metallurgy, 108*, 345-352.

Fan, G., Chen, F., Chen, D., & Dong, Y. (2020). Recognizing multiple types of rocks quickly and accurately based on lightweight CNNs model. *IEEE Access, 8*, 55269-552787.

Ghasemi, E., Ataei, M., Shahriar, K., Sereshki, F., Jalali, S. E., & Ramazanzadeh, A. (2012). Assessment of roof fall risk during retreat mining in room and pillar coal mines. *International Journal of Rock Mechanics & Mining Sciences, 54*, 80-89.

Ghosh, G. K., & Sivakumar, C. (2018). Application of underground microseismic monitoring for ground failure and secure longwall coal mining operation: a case study in an Indian mine. *Journal of Applied Geophysics, 150*, 21-39.

Glorot, X., & Bengio, Y. (2010). Understanding the difficulty of training deep feedforward neural networks. *Proceedings of the Thirteenth International Conference on Artificial Intelligence and Statistics*, (pp. 249-256).

Han, S., Li, H., Li, M., & Luo, X. (2019). Measuring rock surface strength based on spectograms with a deep convolutional networks. *Computers & Geosciences, 133*.

Haneberg, W. C. (2008). Using close range terrestrial digital photogrammetry for 3-d slope modeling and discontinuity mapping the United States. *Bulleting of Engineering Geology and the Environment, 67*, 457-469.

He, K., Zhang, X., Ren, S., & Sun, J. (2016). Deep residual learning for image recognition. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, (pp. 770-778).

He, M., Zhang, Z., Ren, J., Huan, J., Li, G., Chen, Y., & Li, N. (2019). Deep convolutional neural network for fast determination of the rock strength parameters using drilling data. *International Journal of Rock Mechanics and Mining Sciences, 123*.

Huang, H., Li, Q., & Zhang, D. (2018). Deep learning based image recognition for crack and leakage defects of metro shield tunnel. *Tunnelling and Underground Space Technology, 77*, 166-176.

Huang, L., Li, J., Hao, H., & Li, X. (2018). Micro-seismic event detection and location in underground mines by using convolutional neural networks (CNN) and deep learning. *Tunnelling and Underground Space Technology, 81*, 265-276.

Iannacchione, A. T., Bajpayee, T. S., & Edwards, J. L. (2005). Forecasting roof falls with monitoring technologies – a look at the moonee colliery experience. In S. S. Peng, C. Mark, G. L. Finfinger, S. C. Tadolini, K. A. Heasley, & A. W. Khair (Ed.), *Proceedings of the 24th International Conference on Ground Control in Mining*, (pp. 44-51). Morgantown, WV.

Iannacchione, A. T., Dolinar, D. R., & Mucho, T. P. (2002). High stress mining under shallow overburden in underground U.S. stone mines. *Proceedings of the First International Seminar on Deep and High-Stress Mining* (pp. 1-11). Nedlands, Australia: Australian Centre for Geomechanics.

Iannacchione, A. T., Dolinar, D. R., Prosser, L. J., Marshall, T. E., Oyler, D. C., & Compton, C. S. (1998). Controlling roof beam failures from high horizontal stresses in underground stone mines. In S. S. Peng (Ed.), *Proceedings of the 17th International Conference on Ground Control in Mining*, (pp. 102-112). Morgantown, WV.

Iannacchione, A. T., Prosser, L. J., Esterhuizen, G. S., & Bajpayee, T. S. (2006). Assessing roof fall hazards for underground stone mines: a proposed methodology. *2006 SME Annual Meeting and Exhibition* (pp. 1-9). St. Louis, MO: Society for MIning, Metallurgy, and Exploration.

Iannacchione, A. T., Prosser, L. J., Esterhuizen, G. S., & Bajpayee, T. S. (2007). Methods for determining roof fall risk in underground mines. *Mining Engineering, 59*(11), 47-53.

Iannacchione, A., Miller, T., Esterhuizen, G., Slaker, B., Murphy, M., Cope, N., & Thayer, S. (2020). Evaluation of stress-control layout at the Subtropolis Mine, Petersburg, Ohio. *International Journal of Mining Science and Technology, 30*, 77-83.

Jain, A. K., & Farrokhnia, F. (1991). Unsupervised texture segmentation using Gabor filters. *Pattern Recognition, 24*(12), 1167-1186.

Jiang, W., Lam, K. M., & Shen, T. Z. (2009). Efficient edge detection using simplified Gabor wavelenghts. *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics), 39*(4), 1036-1047.

Johnson, J. M., & Khoshgoftaar, T. M. (2020). The effects of data sampling with deep learning and highly imbalanced big data. *Information Systems Frontiers, 22*, 1113-1131.

Keskar, N. S., Mudigere, D., Nocedal, J., Smelyanskiy, M., & Tang, P. T. (2016). On large-batch training for deep learning: generalization gap and sharp minima. *arXiv preprint arXiv: 1609.04836*.

LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. *Nature, 521*, 436-444.

Li, J., Zhang, L., Wu, Z., Ling, Z., Cao, X., Guo, K., & Yan, F. (2020). Autonomous Martian rock image classification based on transfer deep learning methods. *Eath Science Information, 13*, 951-963.

Li, N., Hao, H., Gu, Q., Wang, D., & Hu, X. (2017). A transfer learning method for automatic identification of sandstone microscopic images. *Computers & Geosciences, 103*, 111-121.

Luo, M., Wang, K., Cai, Z., Liu, A., Li, Y., & Cheang, C. F. (2019). Using imbalanced triangle synthetic data for machine learning anomaly detection. *Comput. Mater. Continua, 58*(1), 15-26.

Makridakis, S. (2017). The forthcoming Artificial Intelligence (AI) revolution: its impact on society and firms. *Futures, 90*, 46-60.

Mark, C., & Mucho, T. P. (194). Longwall mine design for control of horizontal stresses. *Proceedings: U.S. Bureau of Mines Technology Transfer Seminar*, (pp. 53-76). Pittsburg, PA.

Mine Safety and Health Administration Part 50 Data. (2009 - 2019). Retrieved from https://www.msha.gov/data-reports/data-sources-calculators

NASA/Goddard Space Flight Center Scientific Visualization Studio. (2004, June 14). *Tectonic Plates and Plate Boundaries (WMS)*. Retrieved May 8, 2020, from https://svs.gsfc.nasa.gov/2953

Neyshabur, B., Sedghi, H., & Zhang, C. (2020). What is being transferred in transfer learning? *Advances in Neural Information Processing Systems, 33*.

Pan, S. J., & Yang, Q. (2010). A survey on transfer learning. *IEEE Transactions on Knowledge and Data Engineering, 22*, 1345-1359.

Pappas, D. M., & Mark, C. (2012). Roof and rib fall Incident trends: a 10-year profile. *Trans Soc Min Metal Explor, 330*, 462-478.

Planche, B., Wu, Z., Ma, K., Sun, S., Kluckner, S., Lehmann, O., Chen, T., Hutter, A., Zakharov, S., Kosch, H., Ernst, J. (2017). Depthsynth: real-time realistic synthetic data generation from cad Models for 2.5d recognition. *2017 International Conference on 3D Vision* , (pp. 1-10). Qingdao.

Prusek, S., Rajwa, S., Wrana, A., & Krzemien, A. (2017). Assessment of roof fall risk in longwall coal mines. *International Journal of Mining, Reclamation and Environment, 31*(8), 558-574.

Qin, C., & Chian, S. C. (2018). Revisiting crown stability of tunnels deeply buried in non-uniform rock surrounds. *Tunnelling and Underground Space Technology, 73*, 154-161.

Selvaraju, R. R., Cogswell, M., Das, A., Vedantam, R., Parikh, D., & Batra, D. (2017). Grad-cam: visual explanations from deep networks via gradient-based localization. *Proceedings of the IEEE International Conference on Computer Vision*, (pp. 618-626).

Sengani, F. (2020). The use of ground penetrating radar to distinguish between seismic and non-seismic hazards in hard rock mining. *Tunnelling and Underground Space Technology, 103*.

Shen, B., Guo, H., King, A., & Wood, M. (2006). An integrated real-time roof monitoring system for underground coal mines. In N. Aziz (Ed.), *Coal 2006: Coal Operators' Conference, University of Wollongong & the Australasian Institute of MIning and Metallurgy*, (pp. 64-76).

Shen, W., Bai, J., Li, W., & Wang, X. (2018). Prediction of relative displacement for entry roof with weak plane under the effect of mining abutment stress. *Tunnelling and Underground Space Technology, 71*, 309-317.

Sherizadeh, T., & Kulatilake, P. H. (2016). Assessment of roof stability in a room and pillar coal mine in the U.S. using three-dimensional distinct element method. *Tunnelling and Underground Space Technology, 59*, 24-37.

Shi, S., Zhao, R., Li, S., Xie, X., Li, L., Zhou, Z., & Liu, H. (2019). Intelligent prediction of surrounding rock deformation of shallow buried highway tunnel and its engineering application. *Tunnelling and Underground Space Technology, 90*, 1-11.

Shrikumar, A., Greenside, P., & Kundaje, A. (2017). Learning important featuers through popagating activation differences. *Proceedings of the 34th International Conference on Machine*, (pp. 3145-3153).

Srivastava, N., Hinton, G., Krizhevsky, A., Sutskever, I., & Salakhutdinov, R. (2014). Dropout: a simple way to prevent neural networks from overfitting. *Journal of Machine Learning Research, 15*, 1929-1958.

Sundararajan, M., Taly, A., & Yan, Q. (2017). Axiomatic attribution for deep networks. *Proceedings of the 34th International Conference on Machine Learning*, (pp. 3319-3328).

TensorFlow. (n.d.). *Integrated Gradients*. Retrieved December 22, 2020, from https://www.tensorflow.org/tutorials/interpretability/integrated_gradients. CC by 4.0.

Tremblay, J., Prakash, A., Acuna, D., Brophy, M., Jampani, V., Anil, C., To, T., Cameracci, E., Boochoon, S., Birchfield, S. (2018). Training deep networks with synthetic data: bridging the reality gap by domain randomization. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR) Workshops*, (pp. 969-977).

Van Grinsven, M. J., Van Ginneken, B., Hoyng, C. B., Theelen, T., & Sanchez, C. I. (2016). Fast convolutional neural network training using selective data sampling: application to hemorrhage detection in color fundus images. *IEEE Transactions on Medical Imaging, 35*(5), 1273-1284.

Wan, Z., Zhang, Y., & He, H. (2017). Variational autoencoder based synthetic data generation for imbalanced learning. *2017 iEEE Symposium on Computational Intelligence.* Honolulu.

Ward, D., Moghadam, P., & Hudson, N. (2018). Deep leaf segmentation using synthetic data. *British Machine Vision Conference.*

Wilkins, A. H., Strange, A., Duan, Y., & Luo, X. (2020). Identifying microseismic events in a mining scenario using a convolutional neural network. *Computers & Geosciences, 137*.

World Economic Forum. (2017). *Digitial Transformation Initiative: Mining and Metals Industry.* World Economic Forum. Retrieved from https://www.accenture.com/_acnmedia/Accenture/Conversion-Assets/WEF/PDF/Accenture-Mining-And-Metals-Industry.pdf

Wu, R., Fujita, Y., & Soga, K. (2020). Integrating domain knowledge with deep learning models: an interpretable AI system for automatic work prgoress identification of NATM tunnels. *Tunneling and Underground Space Technology, 105*.

Xue, Y., Cai, X., Shadabfar, M., Shao, H., & Zhang, S. (2020). Deep learning-based automatic recognition of water leakage area in shield tunnel lining. *Tunneling and Underground Space Technology, 104.*

Yosinski, J., Clune, J., Bengio, Y., & Lipson, H. (n.d.). How transferable are features in deep neural networks? *Advances in Neural Information Processing Systems 27.* NIPS Foundation.

Zhang, D. B., Zhang, Y., Cheng, T., Meng, Y., Fang, K., Garg, A., & Garg, A. (2017). Measurement of displacement for open pit to underground mining transition using digital photogrammetry. *Measurement, 109*, 187-199.

Zhang, J., Tan, T., & Ma, L. (2002). Invariant texture segmentation via circular Gabor filters. *Object recognition supported by user interaction for service robots*, 901-904.

Zhang, Q., Wu, N. Y., & Zhu, S. C. (2018). Interpretable convolutional neural networks. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, (pp. 8827-8836).

Zhao, S., Zhang, D. M., & Huang, H. W. (2020). Deep learning-based image instance segmentation for moisture marks of shield tunnel lining. *Tunnelling and Underground Space Technology, 95*.

Zhao, Y., Zhang, N., & Si, G. (2016). A fiber bragg grating-based monitoring system for roof safety control in underground coal mining. *Sensors, 16*(10).

APPENDIX A

INTEGRATED GRADIENTS OF TEST IMAGES

In this appendix, the results of the integrated gradients analysis on the test dataset is presented. Table A.1 shows the integrated gradients results of true positive images. True positives are the hazardous roof images classified as hazardous. Table A.2 shows the integrated gradients results of true negative images. True negatives are the non-hazardous roof images classified as non-hazardous. Table A.3 shows the integrated gradients results of false positive images. False positives are the non-hazardous roof images classified as hazardous. Table A.4 shows the integrated gradients results of False negative images. False negatives are the hazardous roof images classified as non-hazardous.

Table A.1 Integrated gradients of true positives

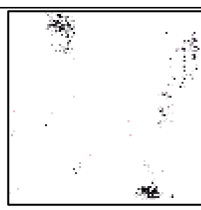
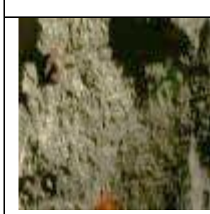
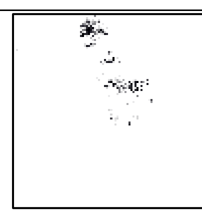
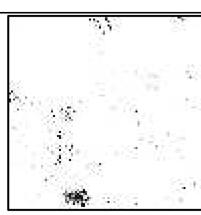
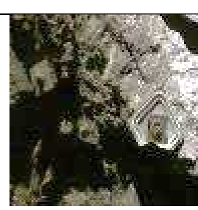
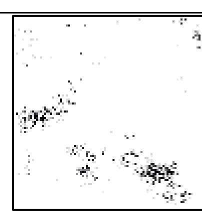
Original	Integrated Gradients	Original	Integrated Gradients
			
			
			

Table A.1 continued

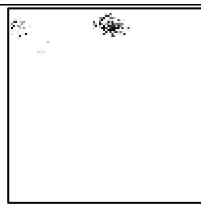
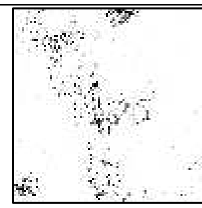
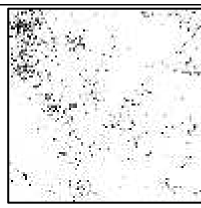
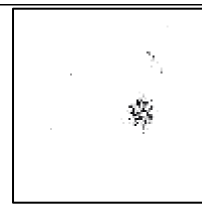
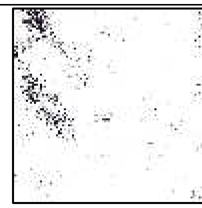
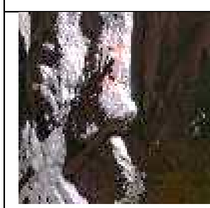
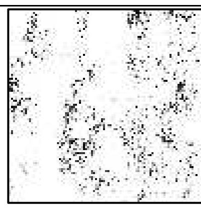
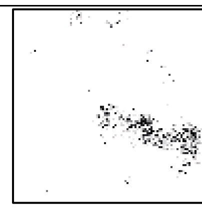
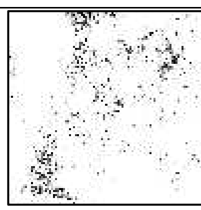
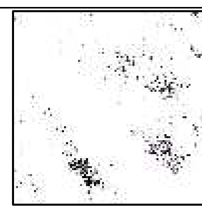
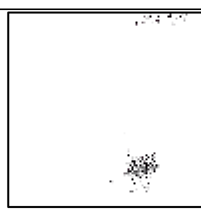
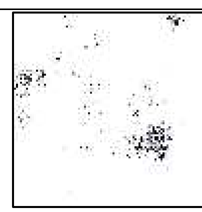
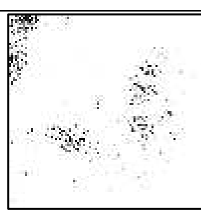
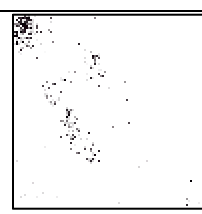
Original	Integrated Gradients	Original	Integrated Gradients
			
			
			
			
			
			
			

Table A.1 continued

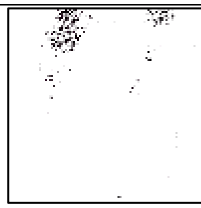
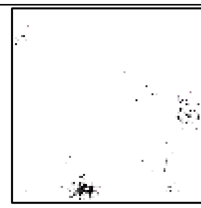
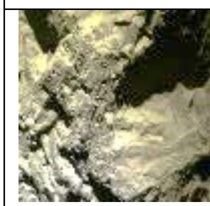
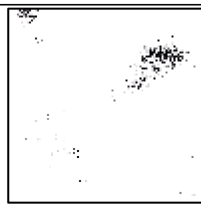
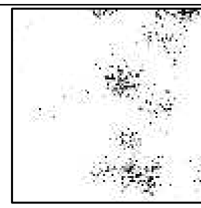
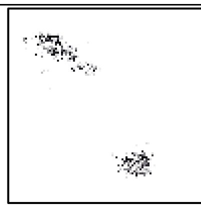
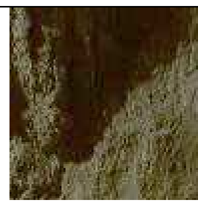
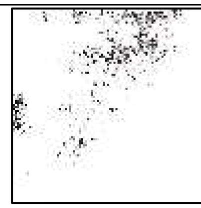
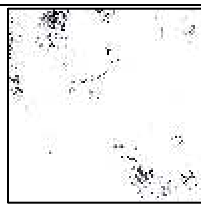
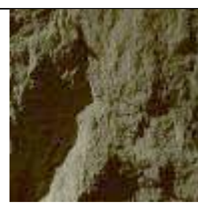
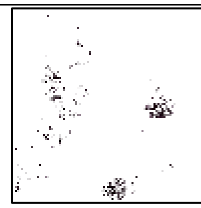
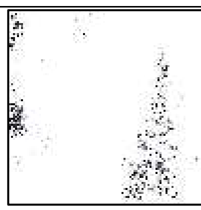
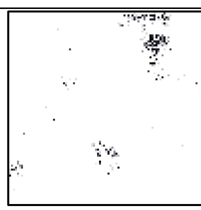
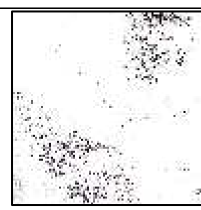
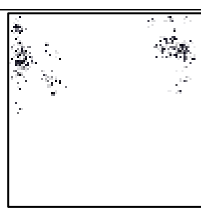
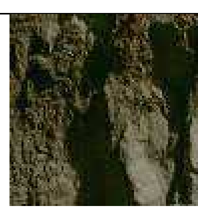
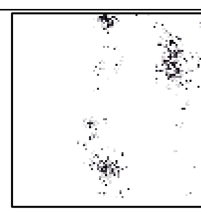
Original	Integrated Gradients	Original	Integrated Gradients
			
			
			
			
			
			
			

Table A.1 continued

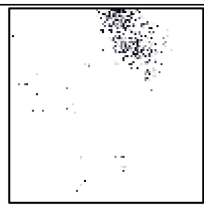
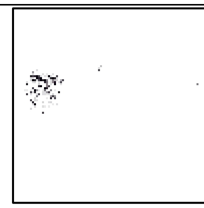
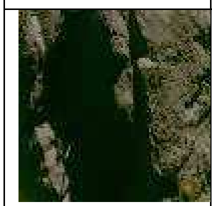
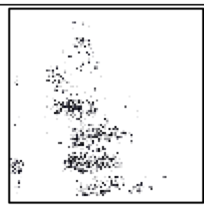
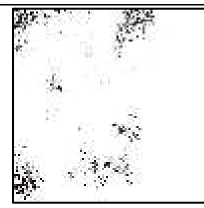
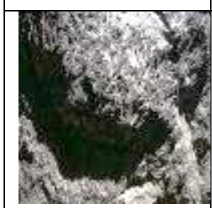
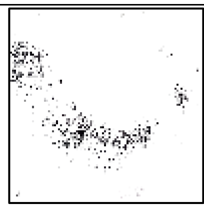
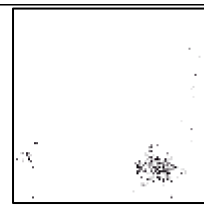
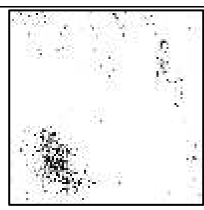
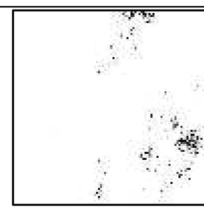
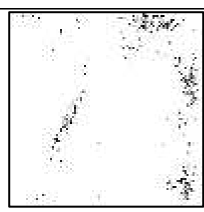
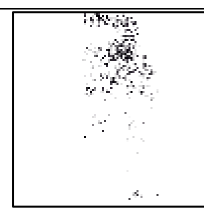
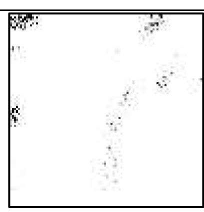
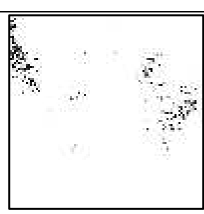
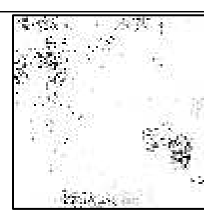
Original	Integrated Gradients	Original	Integrated Gradients
			
			
			
			
			
			
			

Table A.1 continued

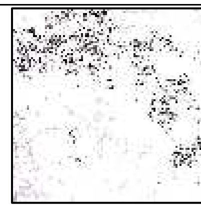
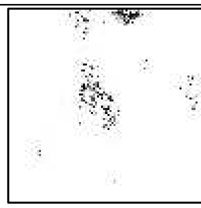
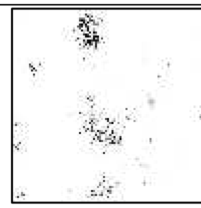
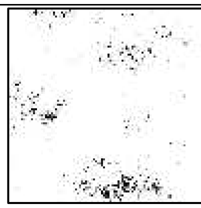
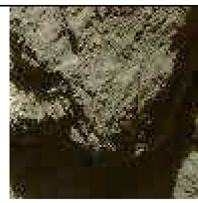
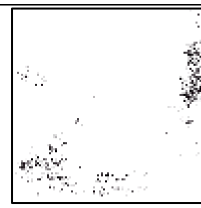
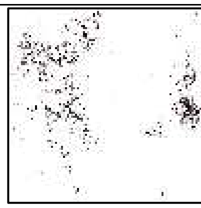
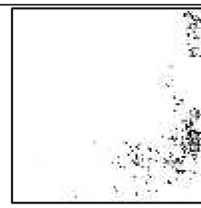
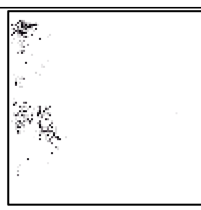
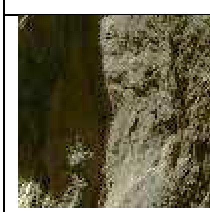
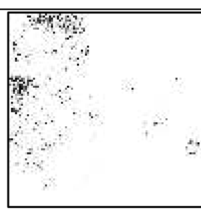
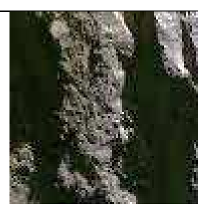
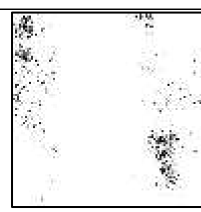
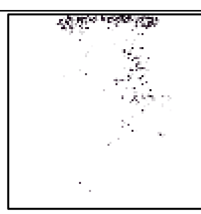
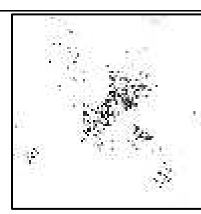
Original	Integrated Gradients	Original	Integrated Gradients
			
			
			
			
			
			
			

Table A.1 continued

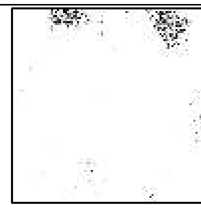
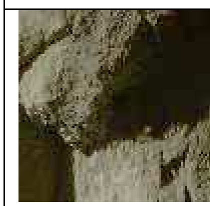
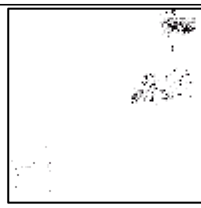
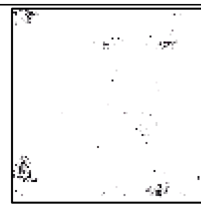
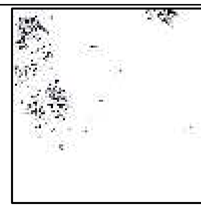
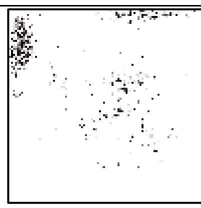
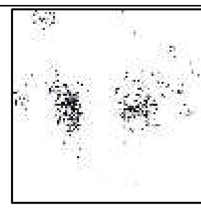
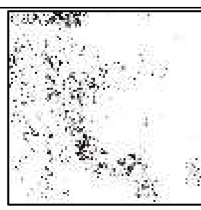
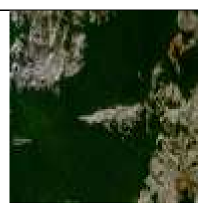
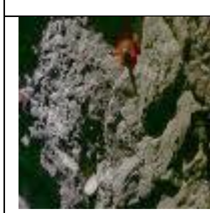
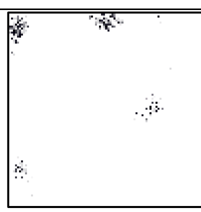
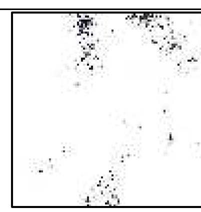
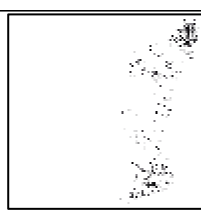
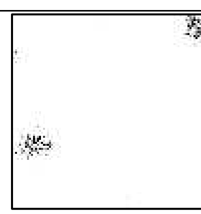
Original	Integrated Gradients	Original	Integrated Gradients
			
			
			
			
			
			
			

Table A.1 continued

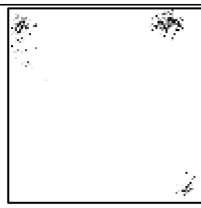
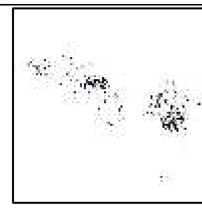
Original	Integrated Gradients	Original	Integrated Gradients
			
			

Table A.2 Integrated gradients of true positives

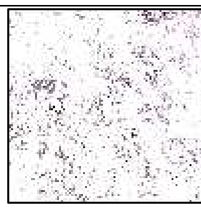
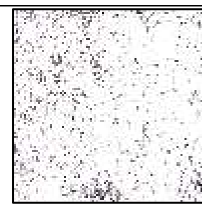
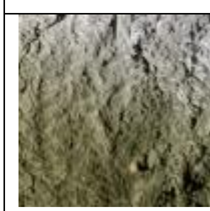
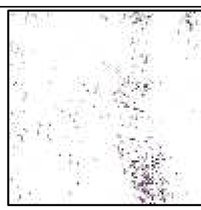
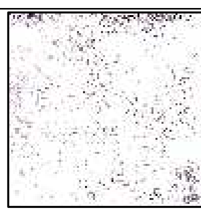
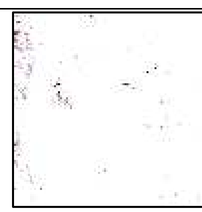
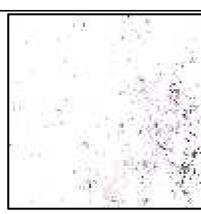
Original	Integrated Gradients	Original	Integrated Gradients
			
			
			
			

Table A.2 continued

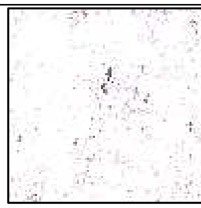
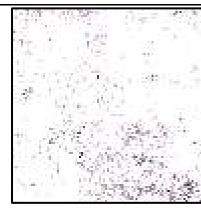
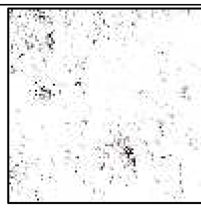
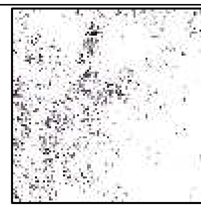
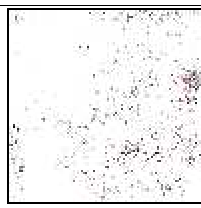
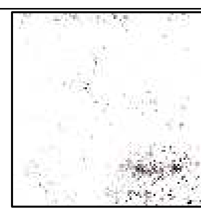
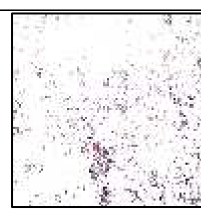
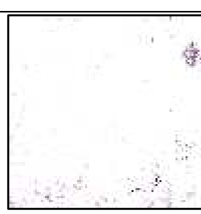
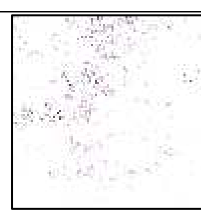
Original	Integrated Gradients	Original	Integrated Gradients
			
			
			
			
			
			
			

Table A.2 continued

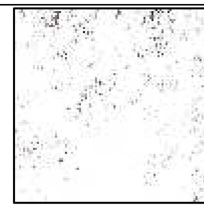
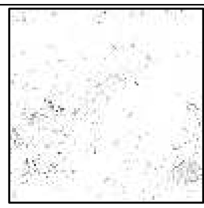
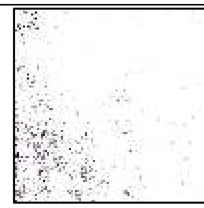
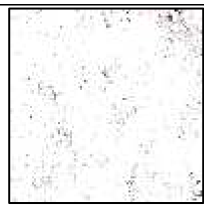
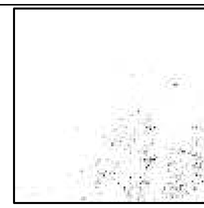
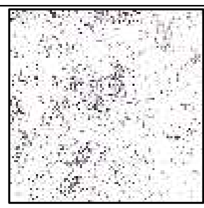
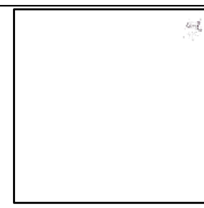
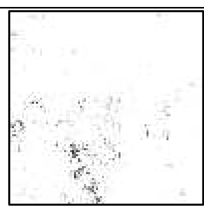
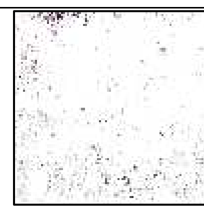
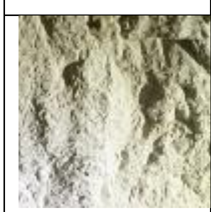
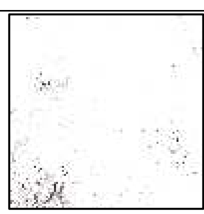
Original	Integrated Gradients	Original	Integrated Gradients
			
			
			
			
			
			
			

Table A.2 continued

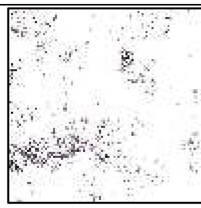
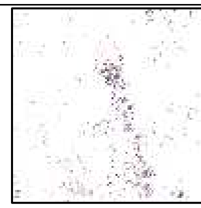
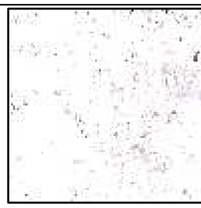
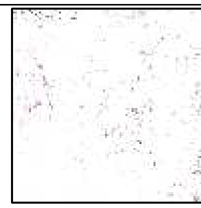
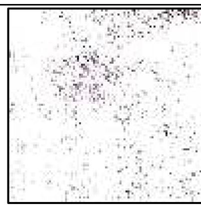
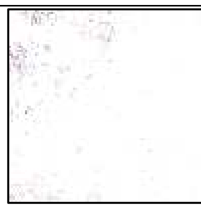
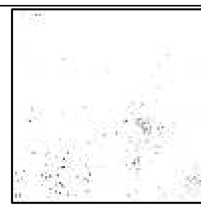
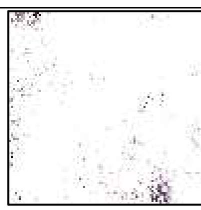
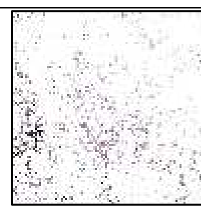
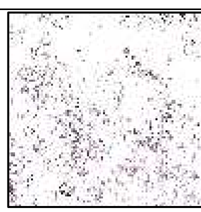
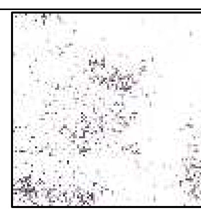
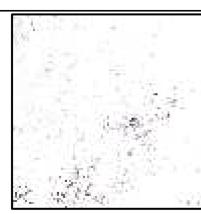
Original	Integrated Gradients	Original	Integrated Gradients
			
			
			
			
			
			
			

Table A.2 continued

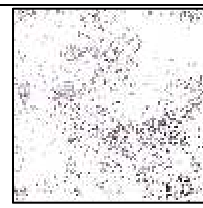
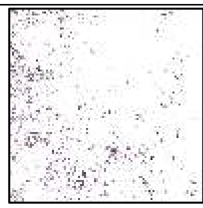
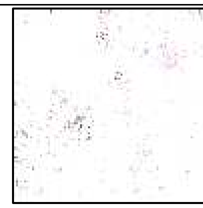
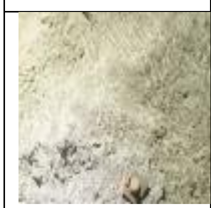
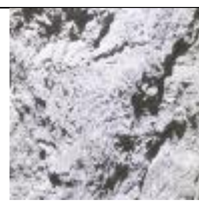
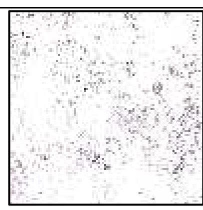
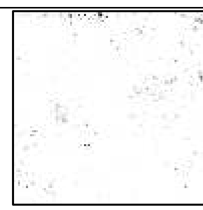
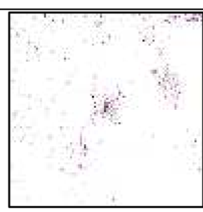
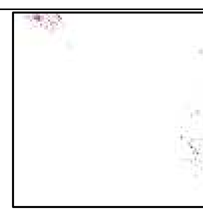
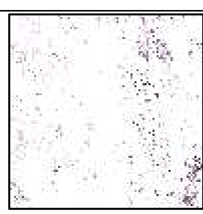
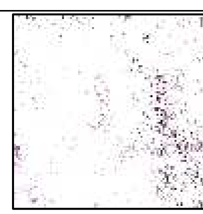
Original	Integrated Gradients	Original	Integrated Gradients
			
			
			
			
			
			
			

Table A.2 continued

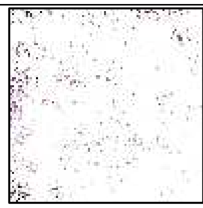
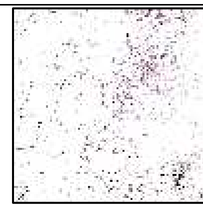
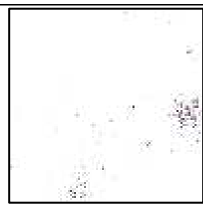
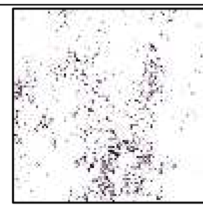
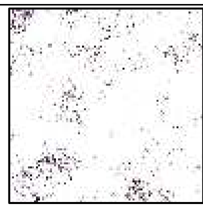
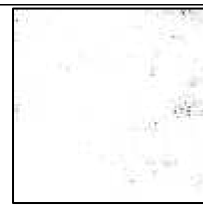
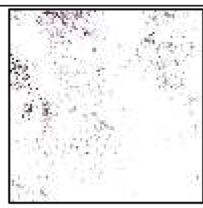
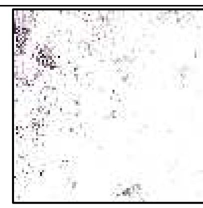
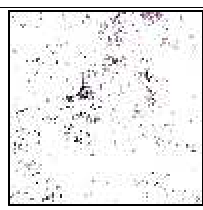
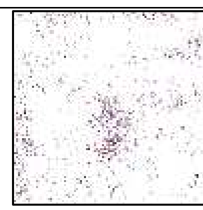
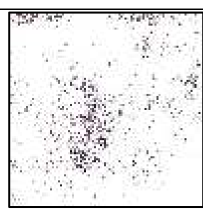
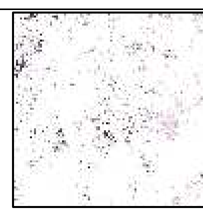
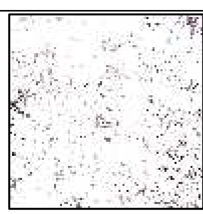
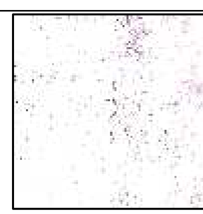
Original	Integrated Gradients	Original	Integrated Gradients
			
			
			
			
			
			
			

Table A.2 continued

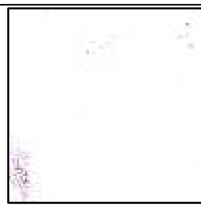
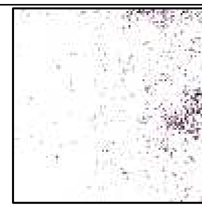
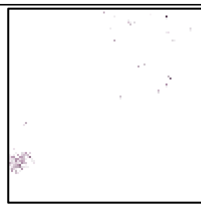
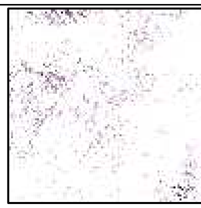
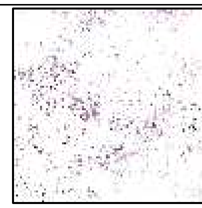
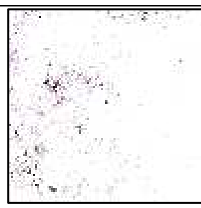
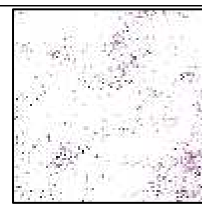
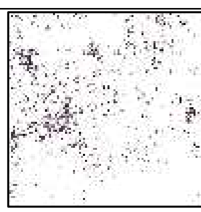
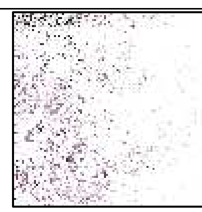
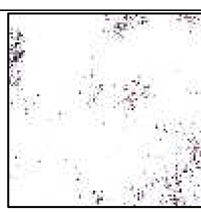
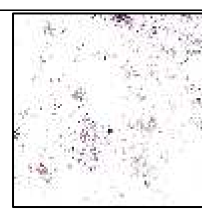
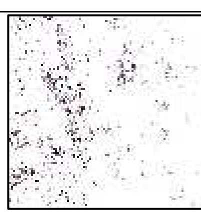
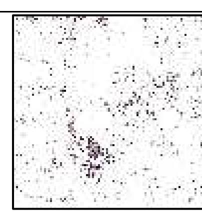
Original	Integrated Gradients	Original	Integrated Gradients
			
			
			
			
			
			
			

Table A.2 continued

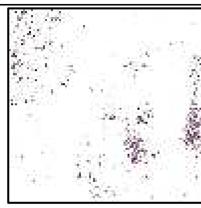
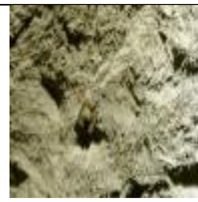
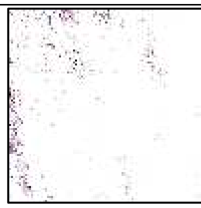
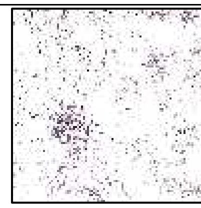
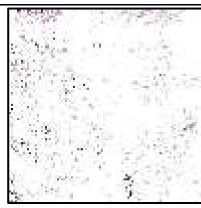
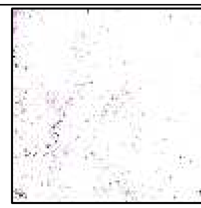
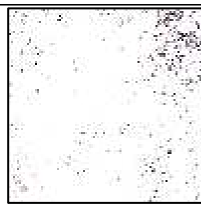
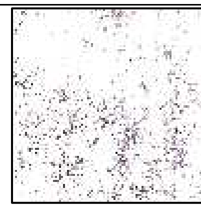
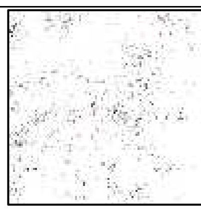
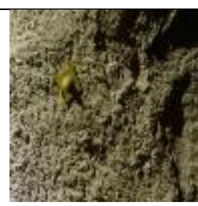
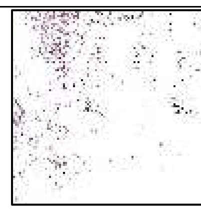
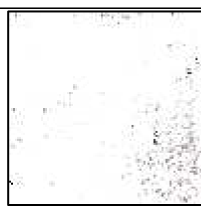
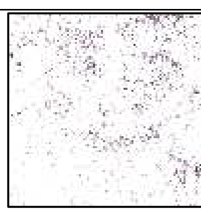
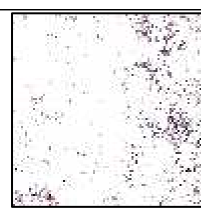
Original	Integrated Gradients	Original	Integrated Gradients
			
			
			
			
			
			
			

Table A.2 continued

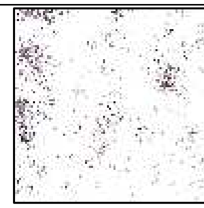
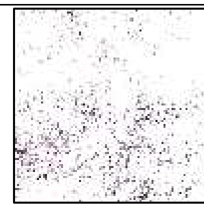
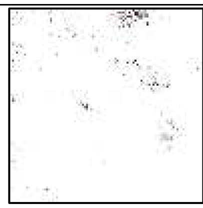
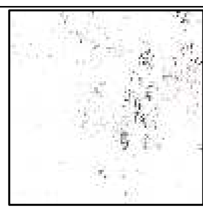
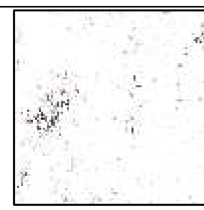
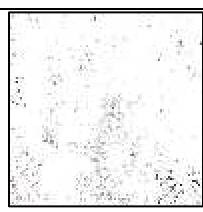
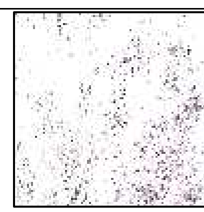
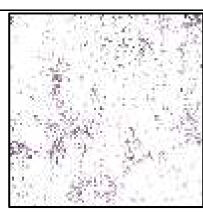
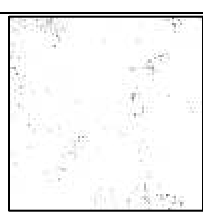
Original	Integrated Gradients	Original	Integrated Gradients
			
			
			
			
			
			
			

Table A.2 continued

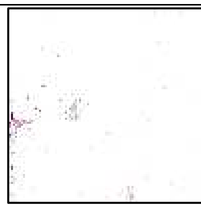
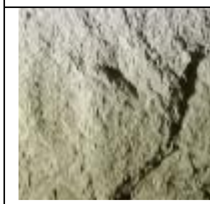
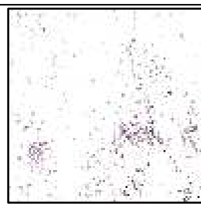
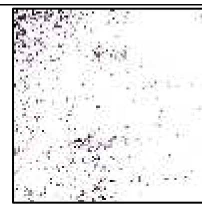
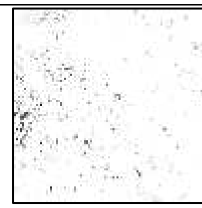
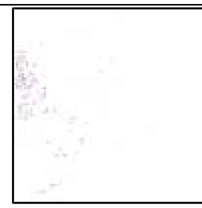
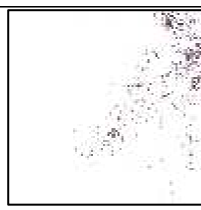
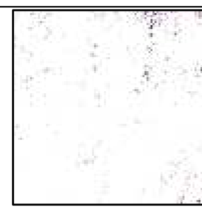
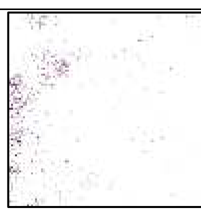
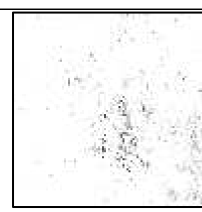
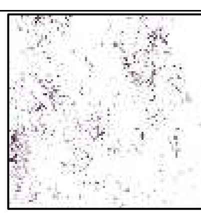
Original	Integrated Gradients	Original	Integrated Gradients
			
			
			
			
			
			
			

Table A.2 continued

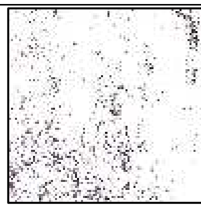
Original	Integrated Gradients	Original	Integrated Gradients
			

Table A.3 Integrated gradients of false positives

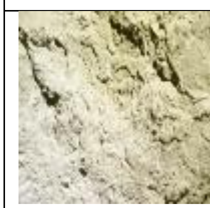
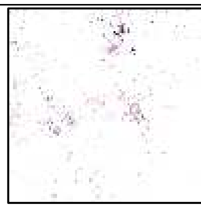
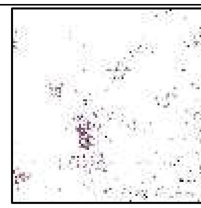
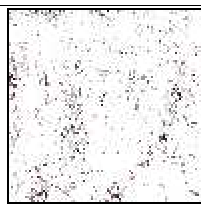
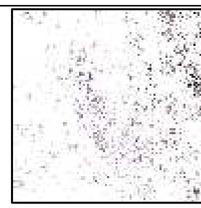
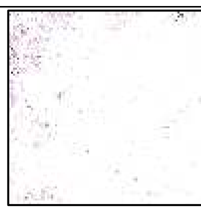
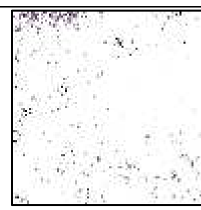
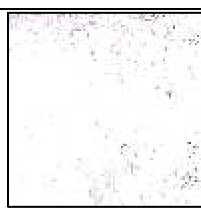
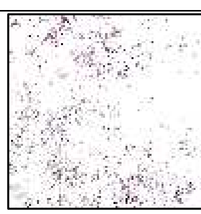
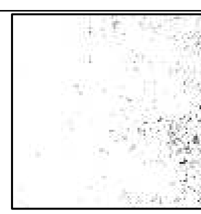
Original	Integrated Gradients	Original	Integrated Gradients
			
			
			
			
			

Table A.3 continued

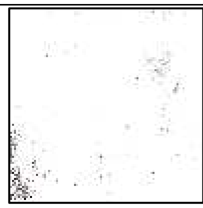
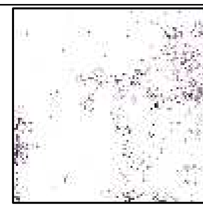
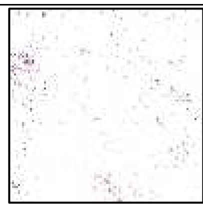
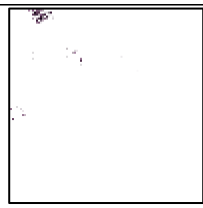
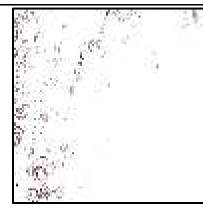
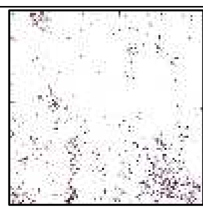
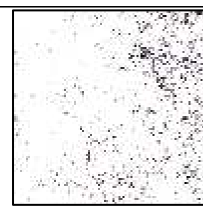
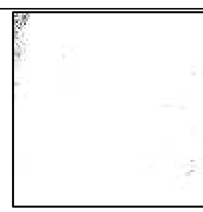
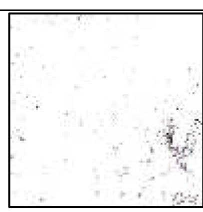
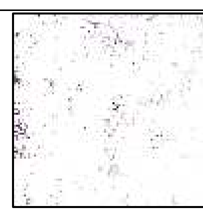
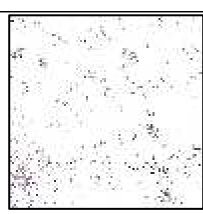
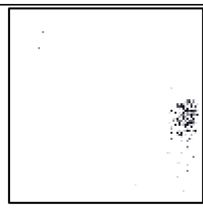
Original	Integrated Gradients	Original	Integrated Gradients
			
			
			
			
			
			
			

Table A.4 Integrated gradients of false positives

Original	Integrated Gradients
	

APPENDIX B

PERMISSION TO USE COPYRIGHTED MATERIALS

Ergin,

On behalf of SME, I authorize the reuse of the figure noted below for your use as described. Please include a full citation as to the source of the image.

Jane Olivier
Manager of Book Publishing
SME

Begin forwarded message:

From: SME Customer Service <cservice@smenet.org>
Subject: FW: Image reproduction request
Date: December 22, 2020 at 8:16:04 AM MST
To: "Jane Olivier (Home)" <glacierpublishing@gmail.com>

Hey Jane, hope you are well, below is a request for one of yours. K

From: Ergin Isleyen [mailto:isleyen@mymail.mines.edu]
Sent: Tuesday, December 22, 2020 7:46 AM
To: SME Customer Service <cservice@smenet.org>
Subject: Re: Image reproduction request

WARNING: This message came from outside of SME. Do not click on links or attachments unless you know the content to be safe.

Thank you for your quick response. I need SME's permission to use Figure 6.4-1 Room and pillar method from SME Mining Engineering Handbook 3rd edition.
Best regards,
-Ergin

Figure B.1 Permission to reuse Figure 2.1 from Darling, 2011