DETECTION OF RAIL FAULTS FROM GRAY-LEVEL IMAGES WITH SEGMENTATION-BASED DEEP NETWORKS

Master's Thesis

Fatih ERDEN

Eskişehir 2024

DETECTION OF RAIL FAULTS FROM GRAY-LEVEL IMAGES WITH SEGMENTATION-BASED DEEP NETWORKS

Fatih ERDEN

Master's Thesis

Department of Electrical and Electronics Engineering

Programme in Electronics

Supervisor: Asst. Prof. Dr. Mehmet FİDAN

(Co-Supervisor: Assoc. Prof. Dr. Ömür AKBAYIR)

Eskişehir

Eskişehir Technical University

Institute of Graduate Programs

July 2024

FINAL APPROVAL FOR THESIS

This thesis titled DETECTION OF RAIL FAULTS FROM GRAY-LEVEL IMAGES WITH SEGMENTATION-BASED DEEP NETWORKS has been prepared and submitted by Fatih ERDEN in partial fulfillment of the requirements in “Eskişehir Technical University Directive on Graduate Education and Examination” for the Degree of Master's in Electrical and Electronics Engineering Department has been examined and approved on 09/07/2024.

Committee Members		**Title, Name and Surname**	**Signature**
Member	:	Asst. Prof. Dr Mehmet FİDAN	
Member	:	Asst. Prof. Dr Nazım İMAL	
Member	:	Asst. Prof. Dr Şener AĞALAR	

Prof. Dr. Semra KURAMA
Director of the Institute of Graduate Programs

09/07/2024

SUPERVISOR APPROVAL

Master's student Fatih ERDEN, whom I supervise, has completed this thesis titled DETECTION OF RAIL FAULTS FROM GRAY-LEVEL IMAGES WITH SEGMENTATION-BASED DEEP NETWORKS . According to my inspections, the work is scientifically and ethically appropriate for the student to the thesis defense exam.

Supervisor
Asst. Prof. Dr. Mehmet FİDAN

ABSTRACT

DETECTION OF RAIL FAULTS FROM GRAY-LEVEL IMAGES WITH SEGMENTATION-BASED DEEP NETWORKS

Fatih ERDEN

Department of Electrical and Electronics Engineering

Programme in Electronics

Eskişehir Technical University, Institute of Graduate Programs, July 2024

Supervisor: Asst. Prof. Dr. Mehmet FİDAN

(Co-Supervisor: Assoc. Prof. Dr. Ömür AKBAYIR)

Railway vehicles are widely used by people and load transportation in the world. In Turkey also with the high speed trains the railway transportation became an important part of our lives. These developments of railway vehicles are also causing the health of railways. It is because the number of vehicles, the cargo weights are increased and, with the new high-speed trains are on the rails.

To obtain the maintenance of the rails, the health of rails must be checked in regular intervals, but it is hard due to the long kilometers of rails and with the traffic, it must be done at night. This cause negligence and mistakes to be overlooked. So, we need to do the controlling with an automatic computer system.

In this project, focused on sleeper and rail problems, which are the two most important components of railway. Models are developed to do detection of cracks, defects in rail and checking the straight alignment of the sleepers by using U-NET deep learning algorithm. Although the trained models were used with insufficient and low-quality images, test results were obtained with an accuracy of 92% or more. Training models with sufficient quality and quantity of data is promising.

Keywords: Railway, Rail, Sleepers alignment, U-NET, Deep learning algorithms.

ÖZET

RAY ARIZALARININ GRİ-SEVİYE RESİMLERDEN KESİMLEME TABANLI DERİN AĞLAR İLE TESPİTİ

Fatih ERDEN

Elektrik ve Elektronik Mühendisliği Anabilim Dalı

Elektronik Bilim Dalı

Eskişehir Teknik Üniversitesi, Lisansüstü Eğitim Enstitüsü, Temmuz 2024

Danışman: Dr. Öğr. Üyesi Mehmet FİDAN

İkinci Danışman: Doç. Dr. Ömür AKBAYIR

Demiryolu araçları dünyada insan ve yük taşımacılığında yaygın olarak kullanılmaktadır. Türkiye'de de hızlı trenlerinde devreye girmesi ile demiryolu taşımacılığı hayatımızın önemli bir parçası haline gelmiştir. Çünkü araç sayısı arttı, yük ağırlıkları da arttı ve yeni hızlı trenlerde raylara çıktılar.

Rayların bakımının yapılabilmesi için belirli aralıklarla rayların sağlamlığının kontrol edilmesi gerekmektedir, ancak rayların kilometrelerce uzun olması ve trafik nedeniyle zor olduğundan gece yapılması gerekmektedir. Bu da ihmallere ve hataların gözden kaçmasına neden olur. Bu yüzden kontrolü otomatik bilgisayar sistemleriyle yapmamız gerekmektedir.

Bu projede demiryolunun en önemli iki bileşeni olan travers ve ray sorunlarına odaklanılmıştır. U-NET derin öğrenme algoritması kullanılarak raydaki çatlakların, kusurların tespiti ve traverslerin düz hizalanmasının kontrol edilmesi için modeller eğitilmiştir. Eğitilen modeller, yetersiz ve düşük kalitedeki resimler ile çalışılmasına karşın %92 ve üzeri doğrulukta test sonuçları elde edilmiştir. Yeterli kalitede ve sayıda veriler ile modellerin eğitimi gelecek vaat etmektedir.

Anahtar Sözcükler: Demiryolu, Ray yüzeyi kusur tespiti, Travers hizalama, U-NET, Derin öğrenme algoritmaları.

ACKNOWLEDGEMENTS

I would like to extend my sincerest thanks to TCDD Railway Research and Technology Center (DATEM) for providing us with the gray level image data. The valuable data you have supplied has played a significant role in advancing our projects and conducting our work more efficiently. Thanks to your support and collaboration, we have made significant strides toward achieving our goals.

I would like to thank my supervisor, Asst. Prof. Dr. Mehmet FİDAN for his guidance and patience throughout the study. Also, I would like to thank my co-supervisor Assoc. Prof. Dr. Ömür AKBAYIR, for his guidance, for the network and important information that he provides for this project. I am thankful for their encouragement and motivation. I would like to express my love and gratitude to my family for their support, and always best wishes.

Fatih ERDEN

09/07/2024

STATEMENT OF COMPLIANCE WITH ETHICAL PRINCIPLES AND RULES

I hereby truthfully declare that this thesis is an original work prepared by me; that I have behaved in accordance with the scientific ethical principles and rules throughout the stages of preparation, data collection, analysis and presentation of my work; that I have cited the sources of all the data and information that could be obtained within the scope of this study, and included these sources in the references section; and that this study has been scanned for plagiarism with “scientific plagiarism detection program” used by Eskişehir Technical University, and that “it does not have any plagiarism” whatsoever. I also declare that, if a case contrary to my declaration is detected in my work at any time, I hereby express my consent to all the ethical and legal consequences that are involved.

Fatih ERDEN

CONTENTS

Page

LIST OF FIGURES

Page

LIST OF IMAGES

Page

GLOSSARY OF SYMBOLS AND ABBREVIATIONS

CNN	:	Convolutional Neural Network
AI	:	Artificial Intelligence
DL	:	Deep Learning
ML	:	Machine Learning
SI	:	The International System of Units
MLP	:	Multilayer Neural Networks
ANN	:	Artificial Neural Networks

1. INTRODUCTION

Railway transportation has been a very important transportation element in the world since the past. In our country, railway transportation was started in the early 1800s to transport mostly cargo. As a result of the use of these huge vehicles, some problems and malfunctions may occur in the railways. Some examples of these are cracks in the rails, ground collapses, and misalignment of the sleepers. It is essential to detect these problems when they are at their minimum before they escalate and to start maintenance and repair operations immediately. In this way, possible loss of life and property can be prevented. Of course, problem detection must be carried out using fast and effective methods. In today’s technological world, problem detection is easier with smarter methods. Briefly, the methodology works as follows. Cameras capture kilometers of data of the rails, and then the data are processed by computer programs with image processing algorithms and anomalies are detected. In this way, human errors and oversight are prevented, and the workload is reduced (Yaman, 2018).

1.1. Some Concepts about Railways

Rails are rodding whose components are iron, carbon, silica, manganese, phosphorus and sulfur, which enable railway vehicles to move easily on the line.

Bolt is called the joints of the track where the rails are connected to each other.

Image 1.1. *Fastening model*

A sleeper is a building material that spreads the loads coming from the rail over a larger surface by ensuring that two parallel rails stand in the same alignment. Sleepers can be divided into four groups: iron, wood, concrete and plastic (Çelik, 2015).

Ballast is the ground consisting of hard and solid stones laid on the platform, filling the gaps between the sleepers and supporting the sleepers. Ballast distributes the loads from the sleeper homogeneously over a larger area on the platform. It keeps the road on a fixed axis and protects the road from mud.

Switch is called a track changing mechanism that allows railway vehicles moving on the rail system track to switch to another track (Çelik, 2015).

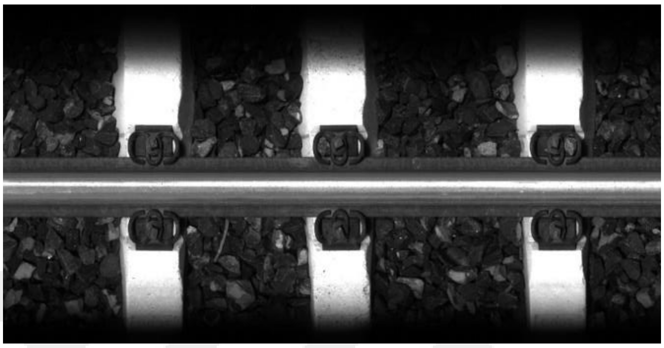


Image 1.2. *Rails, ballast*

1.2. Inspection and Fault Detection

Primitive contact fault detection studies have been caried out by people from past to present but now, in today’s technological world, instead of these methods, the data taken from the rails with the help of cameras and external sensors are processed in the computer environment and rail inspection and fault detection studies are carried out autonomously.

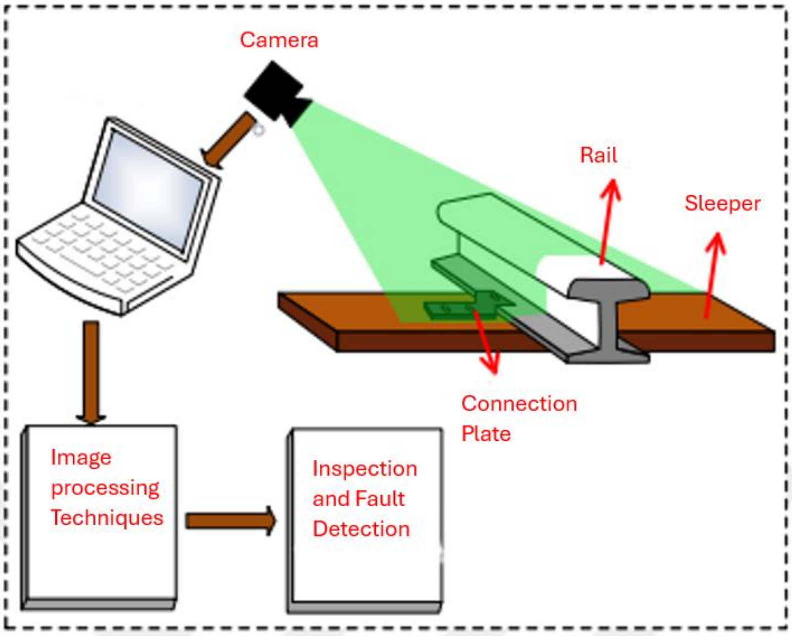


Image 1.3. *Data capture and rail model*

Above picture one can see the model of and railway and the methodology of data verification from the environment (Yaman, 2018).

1.3. Purpose and Scope of the Thesis

In this project, image processing algorithm approaches on deep learning were implemented by working on the data set that is already obtained from railway tracks. The main purpose of these algorithms is to detect cracks in the rail line and alignment of sleepers at correct distance on photographs taken from the cameras in the data set. Deep learning algorithms are built using U-NET architecture. This architecture is a new architecture and is planned to be used in this project due to its good performance in segmentation applications. As a result, the aim is to have higher accuracy and working speed in terms of percentage errors compared to studies in the previous literature (Tatli, 2014).

1.4. Structure of the Thesis

This thesis consists of four chapters including the introduction.

In the introduction, some concepts that make up the railway line are explained, and general introduction is made about condition monitoring and fault detection. Also briefly mentioned railway components and the purpose and scope of the project.

2. RAILWAY DATA AND FAULT VERFICATION

In this part of the project, the methods of how the project progressed are discussed in three sub-sections. These include data collection, fault detection and solution method.

2.1. Data and Data Collection from Railway

2.1.1. What is data?

Data is a vital thing in the field of Deep learning projects. It consists of a set of observations and measurements used to train and mechanize the algorithm. For the machine learning algorithm to work properly, the quality and quantity of data is really important. Data can come in different forms and scopes. For example, numeric, categorical. It can also be captured from spreadsheet APIs, databases. This dataset is utilized by deep learning and machine learning algorithms to comprehend patterns and the link between inputs and outputs, which is then used to tasks like classification and prediction.

Data is divided into two types: *Labeled data and Unlabeled data.*

Labeled data includes a label that the program model is trying to identify, but unlabeled data does not contain a label or a target. In machine learning project the data are divided into training and testing sets. The training sets are used to train the algorithm and afterwards, the algorithm is verified by the test data.

Data pre-processing is really important step in machine learning because it can affect the quality of the algorithm after training with the set of data. This step can include labeling data, filtering normalizing and feature selection of data (Balık, 2022).

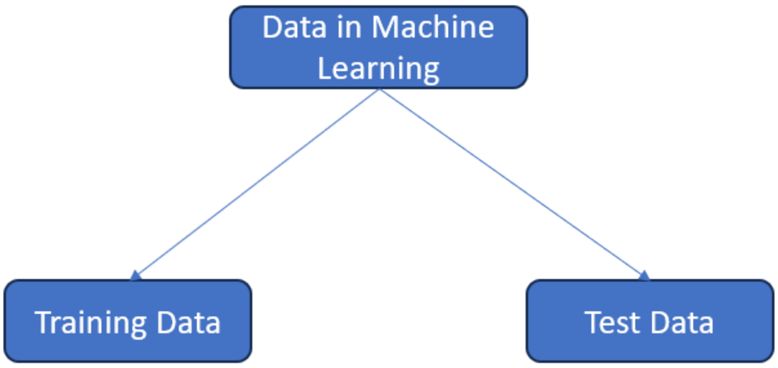
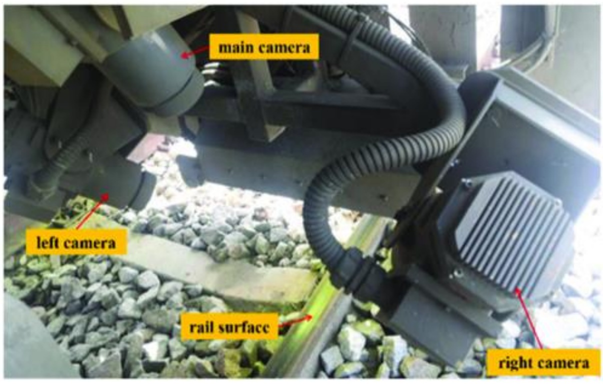


Figure 2.1. *Data classification*

2.1.2. How is the data collected from railways?

The data is gotten from rail ways via and autonomous system. This system has cameras and position sensors that acquires images and 3d images from railways simultaneously. In this project I have gotten these data from TCDD Railway Research and Technology Center (DATEM). Below one can see the representative data acquisition system (Yaman, 2018).



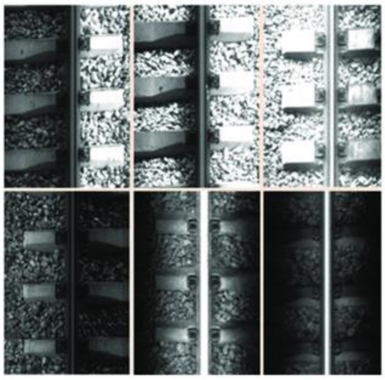


Image 2.1. *Image collection system and example railway images*

2.2. Railway Components and Their Possible Faults

It is important that, in railway transportation, the installation of components that make up the rail lines are very important in terms of safety and comfort. There are many components that make up the railway line available. Such as rail, sleeper, rail connections and ballast, are the most important components to build a railway. These components are shown below image.

Image 2.2. *Components of a railway*

2.2.1. Railway components

Railway transportation is accepted as an important mode of transportation worldwide. Railway parts, which play a role in the functioning of these systems, ensure safe and regular transportation on the lines. Sleepers hold the rails in place while trains move on the tracks. Switches change the trains' route, while connecting materials hold the rails and sleepers together. These components form the solid infrastructure of railway systems and are the basis for safe, fast and efficient transportation. Railway components can be used for both passenger and freight transportation, providing an economical, environmentally friendly and reliable transportation option (Çelik, 2015).

2.2.1.1. Rail

One of the fundamental components of railway systems are rails, which provide the track that trains travel on. Railway lines are supported by rails, which are typically large metal strips that enable trains to move steadily and safely. The movement of the railway wheels on the rails and the contact of the rails with the wheels ensure the smooth operation of the trains.

In the past, steel was preferred when rails were designed and produced during the production stages to ensure their high durability. Consequently, attempts were made to construct railways that would have a long durability from the steel rails, iron joints, concrete sleepers to the smooth surface of the ground.

The infrastructure of railway lines is rails, but correct laying of the rails is required. This ensures safe transport by allowing trains to proceed on the track progressively. The rails are joined end to end with the help of joints at regular intervals and this goes on throughout the railroad tracks.

Railways lines are the basis of rail transport and are extensively used in the sphere of transportation. These make up the fundamentals of secure and proficient rail transport and are very relevant in classes equivalent to public transport, trade and economic growth. Such means can also be considered as one of the types of environmentally friendly transport, as they allow minimizing the negative consequences of road traffic. Of course, if we must make an example, in our country high-speed trains have been seriously affecting the road transport (Yaman, 2018).

2.2.1.2. Sleepers

In railway systems, there exists sleepers whose primary purpose is to enhance the strength as well as durability of the rails, which lies beneath the steel rails. Usually, sleepers which are made of wood, concrete or steel are laid under rails at certain distance between each other. This is done to ensure that the rails have a longer life span and also for increased strength of the rails. It distributes the weight of the trains and the load carried by the rails to ensure that the rails remain stable.

Sleepers are considered to be one of the key parts of the rail infrastructure of paths that is laid in railway lines. Located within the tracks, sleepers help to eliminate undulations in the underlying ground while also supporting the load that rests on the tracks. This way the rails are made stronger, and the railway line is also longer lasting (Yaman, 2018).

2.2.1.3. Ballast

Ballast is the substance that is spread on both sides of the rail track to maintain stability of the rails besides averting the distributing of weights and loads from trains. It largely comprises of stone, gravel and any other related materials. Ballast on the railway is essential structures with effective functions for railway lines. For example, it excludes water and/or mud to accumulate. This disperses pressure all over on the ground and actually makes the ground to become level.

First, ballast is located beneath the rails and assumes the function of fastening the rails. Normal ground is never flat, and the rails are fastened directly to the surface; the ballast then levels the abutments and provides a flat surface that rails rests on. This ensure that the use is safe and less volatile to the rails laid down.

Also, we are to note, it is the ballast which also moderates the vibrations, or, as one may put it, the resonance frequency noticeable upon the passage of trains. This is making both passengers and cargoes more comfortable. For example, ballast is used in systems such as high-speed rail where and it is vital that vibrations are managed, and ballast can provide for that function.

Ballast has an added advantage of offering the facility for draining of water. Water management that direct and drain appropriately under railroad tracks are an effective way that reduce the erosion and improve the life span of rail tracks. Ballast performs functions of impact absorption of water as well as the management of the flow of water by incorporating itself into the drainage system.

Last but not the least, ballast is useful for the management of railway structures Ballast affects the effectiveness of railway lines in supporting railway operations. Ballast also becomes useful when a particular track is destroyed or needs to be changed because of some reasons. During the installation or replacement of rails, the utilization of ballast enables organizations to cut down on working time, and consequently the cost worth.

In its capacity, ballast aggregates provide railway systems with safe, efficient and long endurable transportation solution, upon the occurrence of these factors. An effective ballast provides construction of stable foundations for railway lines, and it is a major rail component in the following ways; supporting the rails, absorbing vibrations, directing water to the ground, and allowing easy access to rail maintenance (Yaman, 2018).

2.2.1.4. Switch

Switches are structures that enable train to staff different tracks. Coilers are generally formed of rolling parts that are connected and located between two rails. Thus, certain aspects of railway traffic are controlled.

Turnouts are also used during maintenance or repair operations on railway lines. When a section on the line to be treated needs to be closed or isolated, the line is closed to train traffic using switches. Other railways continue to operate without interruption.

Switches also play an important role during maintenance or repair operations on railway lines. When a siding on a railway line needs to be closed, switches are used to divert trains to other lines. This allows maintenance or repairs to be carried out and other rail lines to run without interruption.

It is extremely important that switches are designed, installed and operated correctly. Incorrectly positioned or inoperative switches can cause train accidents or serious disruptions to rail traffic. Therefore, railway operators and maintenance teams should take great care to regularly inspect, maintain and, if necessary, repair switches.

As a result, switches are very important components for rail transport. They are only used to control railway traffic by changing the route of trains and allowing them to switch to different lines. Therefore, correct and safe operation of the switches is critical for the safety and efficiency of railway transportation (Yaman, 2018).

2.2.1.5. Connection materials

Fastening materials which are a crucial part of rail systems join rails, sleepers, and other components of the railway staring from a very fundamental structure. These materials ensure that the rails and sleepers are to be fixed and placed correctly as the formation of length/extension of the railway lines. The fixings materials aids in ensure safety and proper running of the rail line.

Fastening materials are components that link the rails; these are of different types. Some of these consist of the bolt, nut, plate, clamp and pendant. These materials make certain that the rails are at the correct position and do not move from this position. They also guarantee the follow-up of the implementation of railway tracks.

This is do by incorporating correcting and fixing material into the railway line construction. The railway can rise above the years with assembled as well as maintained fasteners. This causes the rails and sleepers to be firmly anchored and guarantee that their durability is going to be for a very long time.

Some potential failures are avoided since there are regular checks of connection materials located on railway tracks. Some of the practical examples are bolt tightening, replacement of the plate and the use of clamping. Also, maintenance responsibilities are always met to guarantee secure and sanitary maintenance of the railway line.

Moreover, care for connection materials also enables the use of railway lines for many years without any problems. For example, tightening bolts, replacing plates, and clamps of the dog leg section that is used to connect the high-pressure steam pipes. Further, it contributes to the proper maintenance hence ensuring that the railway line is available for use without intermission. hence, connection materials are one of the railway systems Since they join the rails and sleepers in order to avoid the probable accidents occurring on the railway lines (Yaman, 2018).

2.2.2. Railway faults

2.2.2.1. Rails

Track failures can have serious consequences for the safety and operation of railway lines. Cracks, breaks or deformations that can occur on the rails can stop the journey of trains and even cause accidents. Especially on high-speed rail lines, even the slightest track failure can lead to major hazards. The main causes of track failures are usually factors such as wear, fatigue or exposure to external factors. Timely detection and repair of such failures is vital for the safety of railway lines (Yaman, 2018).

2.2.2.2. Sleepers

Sleepers are one of the basic structures that form the infrastructure of railway lines and are placed under the rails. Decay, breakage, or poor maintenance of sleepers can cause serious problems on railway lines. Defective rails can destabilize the rails, jeopardizing the healthy progress of trains and even causing complete derailment. This could halt the use of the railway line and cause serious disruptions to the transport network (Yaman, 2018).

2.2.2.3. Switches

Switches are mechanisms used on railway lines to switch trains to different lines. Failure of switches negatively affects train traffic. Malfunctions of switch mechanisms such as locking, jamming or breaking cause difficulties in directing trains and completely stop traffic. This can cause delays in train services, congestion on lines and even accidents. Fast detection and repair of switch faults is important to ensure the regular and healthy operation of railway traffic (Yaman, 2018).

2.2.2.4. Connection materials

Fasteners join the railway track, binding the rails with the sleepers. Failures in fasteners can result in the dislodging of rails, the movements of sleepers, or worst-case rope breaks. This might endanger the safe travel of trains and is, therefore, bound to have adverse effects on rail traffic. Regular inspection and proper maintenance of fasteners can be of great effect on the safety and durability of railway lames (Yaman, 2018).

2.2.2.5. Signaling and electrical systems

Failures to the signaling or electrical systems on railway lines may have effects on the safety and punctuality of the trains. Signal light failure or failure on the signaling system may be a problem while routing and speed governing of the trains. Also, failures of electrical systems will halt either the movement or stoppage of trains, thus affecting the traffic on railway lines. Detection of such faults on time and repairing is important in ensuring the safe and efficient operation of railway lines (Yaman, 2018).

3. DEEP LEARNING

In many sectors, so many numbers of projects are being worked with machine learning and artificial intelligence. There is a fundamental reason for this change. It is because, deep learning is a useful and accurate technique used by artificial intelligence to learn complex structures and recognize patterns in complex data sets.

The mathematical models used for deep learning are artificial neural networks. These neural networks consist of layers. Each layer connects to models based on the learning function and performs a specific operation.

Deep learning requires training models that use a large amount of data. This data can be audio, image, text, or various other types. The data for this project are railway images. A deep learning model is trained to recognize objective patterns in images.

Therefore, deep learning is an effective technique that plays an important role in the field of artificial intelligence. It is becoming fashionable, and its importance is increasing day by day. In the future, deep learning methods are expected to become more widespread and will have a wider range of applications (Balcıoğlu, 2023).

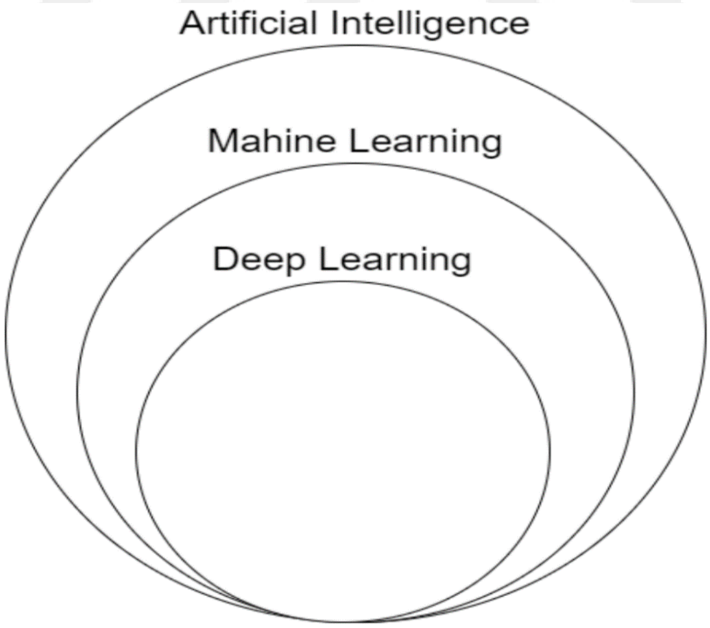


Image 3.1. *Neural network learning*

3.1. Machine Learning

Machine learning gives computer systems the ability to perform specific tasks without being explicitly programmed. Machine learning automates the process of data learning and decision making. This process occurs through looking at pictures and making predictions from those pictures when working on various datasets. In 1950’s and 1960’s are considered to be the years when the basic principles of machine learning were laid. In this period, Alan TURING proposed to measure a machine’s ability to think like human with the “Turing Test”. Arthur Samuel used the term “machine learning” for the first time by developing a computer program that plays checkers. The ‘perceptron’ algorithm was developed by Frank Rosenblatt in the 1960s.It was the first computer model that formed the basis of artificial neural networks. It halted intelligence research, due to the limited capabilities of these models and the problems of multilayer design. In the 1980s, neural networks regained popularity with the development of the backpropagation algorithm. Scientists such as Geoffrey Hinton, Yann LeCun and Yoshua Bengio have done important work on deep learning and advanced neural networks. This work has expanded the application areas of machine learning, especially as the availability of large datasets and powerful processors has increased (Kartal, 2022).

Important Studies and Scientists

1. Geoffrey Hinton: One of the pioneers of deep learning, Hinton popularized the back-propagation algorithm in the 1980s. He also demonstrated the effectiveness of deep learning models by working on “Deep Belief Networks” in 2006.
2. Yann LeCun: LeCun has made significant contributions, especially in the fields of image recognition and computer vision. His 1989 LeNet-5 revolutionized the recognition of handwritten digits and laid the foundation for many deep learning models used today.
3. Andrew Ng: A professor at Stanford University, Ng co-founded the online education platform Coursera and was instrumental in launching the Google Brain project. Ng's work focuses on practical applications of deep learning and big data analytics.
4. Ian Goodfellow: He invented the algorithm known as Generative Adversarial Networks (GANs), which allows two models to compete against each other to produce better results. GANs have enabled groundbreaking developments in many fields such as image and video production.

Machine learning is used in many applications. Risk management and algorithmic trading in the finance sector, disease diagnosis and drug discovery, and autonomous vehicles are examples of the uses of this technology.

Machine learning is also being applied to natural language processing (NLP), speech recognition and recommender systems. The potential of machine learning is even greater, coupled with technologies such as cloud computing and IoT (Internet of Things), think big data analytics.

It has played a central role, and as a result, machine learning can be found in the development of computer science and artificial intelligence. Today, it continues to develop rapidly and transforms our lives. This field has been shaped by the contributions of many important scientists.

There are four main approaches in this area: supervised, unsupervised, reinforcement and semi-supervised learning. Each approach offers different strategies and areas of application depending on the types of data and problems.

3.1.1. Supervised learning

The most popular way of machine learning which predominantly utilized is supervised learning. Supervised learning basically needs training in the model with labelled data. The training data needs to replicate paired sets of features, i.e. specific inputs along with its corresponding accurate outputs as labels. The model will learn the relationship between these paired sets so that it can provide accurate prediction of new unlabeled data later.

Supervised learning can be categorized into two main classes that is classification and regression. To sort a dataset into a category, it becomes a classification problem. An email filtering system can classify an email into a ‘spam’ category or a ‘not spam’ category. Preventing spam is a classification problem. Another example of a classification problem is how a healthcare company can predict whether a patient has a particular disease through certain classification algorithms. Unlike classification problems, regression problems have continuous output. A simple example of regression problems includes predicting a stock price. A more complicated real-world example would be the estimation of house prices (Şeker, Diri, & Balık, 2017).

3.1.2. Unsupervised learning

A method of performing machine learning in which a model is trained from data that do not have specific labels is called unsupervised learning. This method permits the model to attempt to look for structures, patterns, and relative elements in the dataset. Clustering and dimensionality reduction are two vast categories of problems that fall under the shed of unsupervised learning.

Clustering is a group of data whereby data is grouped in a particular way that they are similar do to. For instance, it is possible to employ clustering algorithms where it is possible to cluster the customers based on the frequent shopping behavior. This can be useful in the development of marketing strategies that are business specific and effective. Clustering of gene expression data in different categories with similar genes helps in gaining better focus and dissecting of biological processes in biological data analysis.

Data pre-processing helps prevent machine learning models from training on redundant or unnecessary attributes in immense datasets. Some of the kinds of applications include data visualization as well as feature selection. PCA is used for dimensionality reduction of images, which helps in data exploration and understanding by reducing large datasets into fewer dimensions (Şeker, Diri, & Balık, 2017).

3.1.3. Reinforcement learning

Enhanced learning is the type of learning an agent gains in a given environment either through a reward or through punishment. This one is a strategy that an agent employs in an effort to optimize its actions in an attempt to attain a particular goal. For instance, in the fields of games, robotics and dynamical systems reinforcement learning is popular.

In gaming reinforcement learning algorithms have been employed to develop models that are capable of effectively playing games such as chess or Go. For instance, DeepMind constructed a program called AlphaGo which utilizes reinforcement learning and triumphed over the world champion in Go. Self-driving cars employ a reinforcement learning model to navigate safely by eliminating potential hazards in the surrounding area. In the same way, industrial robots can be taught so as to get the best results in production lines.

3.1.4. Semi-supervised learning

A semi-supervised learning algorithm utilizes both labeled and unlabeled data in order to train a model. This is applied when there is a scarcity of labeled data and requires a combination of the unsupervised and supervised models. It is widely applied in fields like natural language processing NLP and image recognition among others.

Most text data which are used in order to build the language models are in large amounts but unlabeled. These models enable you to not only understand the meaning of the texts but also the context within which the words are used while performing activities such as sentiment analysis or text classification, for example. Semi-supervised learning is specifically applied to the field of image recognition, where there is a large volume of unlabeled images needed to construct machine learning models for specific tasks such as face recognition or object detection. For instance, the Google Photos applies semi-supervised learning to automatically sort and recognize the users' images.

These four fundamental techniques of machine learning are quite useful in many different procedures including problem solving. Each of them has approaches towards different types of data and problems whereas some are more complicated than others. Unsupervised learning: This involves the analysis of unlabeled data with the aim of identifying hidden patterns, while supervised learning: This type of learning involves the use of labeled data to arrive at the most appropriate results. Semi-supervised learning uses labeled and unlabeled data; on the other hand, enhanced learning is one of the learning techniques that boost learning by using the positive and negative reinforcement.

Finally, the significance of Machine learning cannot be undermined as it has been proved immensely beneficial in the advancement of artificial intelligence as well as computer science. This field that grew throughout the years with the help of many great minds is constantly evolving and can already be seen as making a significant impact in our daily lives. The principal approaches to machine learning will underpin the basic techniques in data analysis, predictions, and decision-making in future applications.

3.2. Artificial Neural Network

ANNs are modeled on how the human brain works with an intention of creating computer programs that can act intelligently. These networks are comprised of artificial neurons, consisting of a very large number of simple computational components. These neurons act as input and output interfaces; they receive input signals, transforming them into results which they put out as output signals. Neural networks are usually composed of an input layer, intermediate layer or layers and an output layer.

In 1943, Warren McCulloch and Walter Pitts provided the initial inspiration for artificial neural networks. McMaster and co-workers demonstrated that neurons could be modelled by simple equations. In 1958, Rosenblatt described learning in neural networks by using the perceptron model. In the case of multilayer networks, the backpropagation algorithm devised by Geoffrey Hinton and others in the 1980s was a major step in the training process. This led to the formation of the deep learning category, making it possible for neural networks to perform better in the higher-level tasks.

The following picture illustrates one of the simplest artificial neural network models: Here we have the first layer, second layer, and the third layer which is the output layer (Balcıoğlu, 2023).

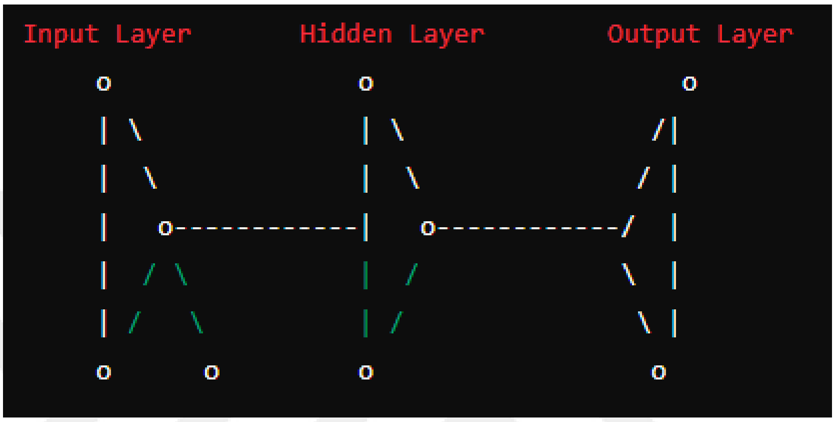


Image 3.2. *Basic artificial network*

It contains an Input Layer where the data acquired by the network is presented. Every neuron has its input feature. Hidden Layer enables computations that map the inputs in a more sophisticated manner to be performed. The network becomes richer and more productive with the development of more than one hidden layer. The last or final layer of the network is the Output Layer, and this layer is also responsible for generating the predicted values of the network.

This example is an artificial neural network which is a computational model. The two are more complicated and powerful due to the fact that real neural networks have more layers and neurons. In the current world, neural networks can be applied in image recognition, speech recognition, natural language processing, among other applications (Balcıoğlu, 2023).

3.3. Convolutional Neural Network

CNNs are deep learning models that are specifically designed for multi-layered data processing, particularly for images and video. Computer vision has evolved from techniques of visual data analysis through the use of CNNs which are very efficient in handling many computer vision tasks. They have their architectural layout of filtering, non-linear transformation, pooling, and density beginning with convolutional and followed by pooling then convolutional and finally a density layer. Convolution layers identify objects that are present within limited region of an image, such as corners, edges, etc., and pooling layers help decrease the size of data in order to lessen the computational load. This is done using fully connected layers which are the final layers of the regression or classification.

These include CNNs are computer vision models that are capable of solving several problems such as image recognition, object detection, and segmentation, among others. However, instead of having to design these features by hand as in conventional machine learning algorithms, CNNs learn them on their own. This has made them more prominent in the field of computing and hence CNNs. CNNs particularly with deep learning technology have been applied in automotive industry, security, medical and many more fields. For instance, applying CNNs in the context of medicine, the devices for the analysis of medical images and disease diagnosis may provide more accurate outcomes with less time consumed (Kaya, 2023).

Advantages

1: Automatic Feature Extraction: CNNs automatically learn important features from the data, eliminating the need for manual feature engineering.

2. Local Connections: Convolution layers learn features of local regions to preserve spatial hierarchy in images.

3. Parameter Sharing: Using the same filters in different image regions reduces the parameters of the model and speeds up the training process.

4. Computational Efficiency: Pooling layers makes data more efficient and reduces overfitting (Alkan, 2023).

Relation with U-NET:

U-NET is a CNN network that has been designed particularly for segmenting images of the medical field. Originally developed by Olaf Ronneberger and colleagues in 2015 and based on an encoder-decoder architecture, U-NET is named for its U-like architecture. In the same way as in the traditional CNNs, the encoder’s function is to extract features, while the decoder aims at generating segmentation masks based on these features.

The connections between the encoder and decoder sections within U-NET are one of its strongest aspects. These bridges enhance the segmentation outcome because the decoder layers receive feature maps with contextual information from encoder layers. This involves both features that are low-level, which means that they encompass fine details, as well as high-level features that encompass meaning.

U-NET extracts key components from CNNs and links them with CNNs to make them optimal for segmentation tasks. On account of the excellence of feature extraction and local connectivity of the CNNs, very high precision segmentation results can be achieved by the U-NET in tasks. This is very important in medical image analysis.

In conclusion, convolutional neural networks have created a significant change in image and video data analysis and can now be applied in many fields. With the benefit of local connections, shared parameters, computational efficiency, and feature extraction on its own, CNN is very good at computer vision problems. U-NET has a special network architecture that optimizes the strengths of CNNs on the task of segmentation. Again, the results obtained above show the potentials and power that a U-NET CNN possesses in fields of this dimension, especially medical image segmentation (Kartal, 2022).

3.4. Segmentation

Image segmentation refers to the process of breaking an image into parts with similar attributes or features. This technique is normally used in locating and identifying objects or regions within the image. Image segmentation is a partition of pixels into different classes wherein each class is considered to represent an object or part of the background in an image. This has wide usage in areas of image processing and computer vision.

Basically, two broad classes of image segmentation are semantic segmentation and instance segmentation. On the other hand, semantic segmentation is a task attributing each pixel in an image to a particular class. All pixels belonging to the same category are now accumulated in that single segment. For example, things that can be grouped under a city image are buildings, roads, cars, and trees. These classes are then detected using semantic segmentation, which segments the image into parts based on these categories. This approach does not make it easy to tell one object of a given class from the other, but instead groups regions in terms of the class they belong to.

While the semantic segmentation can differentiate things belonging to the same class, instance segmentation differentiates everything. For example, if there are multiple people and dogs in a picture, instance segmentation is able to split them and recognize each of them separately. This method portrays each thing as a separate object. Panoptic segmentation is a type of segmentation that is a combination of both semantic and instance segmentation. It divides both the classes of objects and each object that belongs to the classes into groups. This should allow for every pixel in the image to be correctly labelled and each object in the image to be detected.

Applications of image segmentation include areas such as medical imaging, self-driving cars, security and surveillance, manufacturing industry application, and satellite and aerial image analysis. In the medical field, it is employed to detect tumors or organs or other significant structures in medical images such as MR and CT scans. This assists physicians in identifying patients' conditions and developing the proper course of action. In self-driven cars, it is used to detect objects on the road, some of which include other vehicles, people and road signs. This enables a safe and smooth maneuvering of the vehicle.

Land use and cover delineation of satellite imagery involves the use of segmentation techniques. This can be useful for applications such as in urban planning, agricultural and some environmental studies. It becomes vital in security systems as it is responsible for detecting and identifying moving objects. This aids in identifying different extraneous movements in photographs taken in security cameras. This is applied in the production lines for quality checking of the products manufactured and detection of defects in industrial applications. This plays a part in making the production process more efficient and effective.

Image segmentation is generally very exciting, important and useful for different purposes. It enables one to dissect objects, and small elements in an image, with precision. It is used in classifying objects in photos such that there is no need for one to do it his or her own, which is more efficient. It enhances value and effectiveness of operations for instance in industries and the medical field. More accurately and meaningful quantitative results are obtained from image processing. In conclusion, the paper presented showed that image segmentation is beneficial in many areas. This is a very important for the determination and classification of objects and regions and for analysis. It is effective for various applications because it relies on upgraded algorithms and artificial intelligence models that offer high accuracy and performance (Balcıoğlu, 2023).

3.5. What is Deep Network Analysis?

Deep network analysis is a discipline that focuses on the increasing complexity and impact of artificial neural networks in the field of deep learning. Deep learning is a selfcontained branch of methods for machine learning that are used in the analysis of complex data structures. Deep network analysis includes a variety of techniques and methods for understanding how this deep model works, the interactions between inter-layers, and properties. These analyses can be used to improve the performance of deep learning models, avoid overlearning, and discover meaningful patterns, assessing the reliability of the models. Deep network analysis often concerns large, complex data sets, both because of the high amount of data on which deep networks are trained and, inversely, their complex structure. Development and implementation of deep learning models have been an area of research and application, notably, in artificial intelligence, image processing, and natural language processing. Deep network analysis is therefore very basic in better understanding these models and making them more effective in future applications (Zhuang, Qi, & Zhang, 2022).

3.6. Why did we choose Deep Network Analysis with U-NET?

U-NET is applied to a wide range in image segmentation, and in the field of deep learning, it has proved to be very efficient and reliable. It was successfully applied within the framework of this approach, popular in the analysis of medical images. Another good reason why U-NET shall be used is that when compared with a fully connected network, FCN, it is efficient, and besides, it has a low requirement for data training. It is also quite suitable to be used with small amounts of data, since it has been designed in the context of the U-NET architecture.

Another important benefit of the U-NET is that it has an "encoder-decoder" structure. Many convolutional layers, which define the features of the image and down-sample them, compose the encoder submodule. This makes a transition in the image's feature space which gives the true image representation. The decoder part comprises of multiple convolution inverse layers that correspond to the feature map of the input and produces outputs such as classification or segmentation respectively. This structure is effective when used in computing high-resolution results with respect to the segmentation function.

It is also important that U-NET includes a type of connection called jumpers for connecting two devices that are on top of each other. These connections convey the output of the encoder section straight to the decoder section. Since the aspects are learned hierarchically, this helps the model to perform better segmentation by integrating lower-level features with higher level features. This is useful since it permits the model to encompass the whole context but still retain precise structures.

From the above-discussed features of U-NET, this has contributed greatly to fields like medical imaging. The segmentation of medical images with great precision is where deep learning models like U-NET come in handy. These models are beneficial to the doctors and researchers in that they offer them information on how to diagnose the diseases and develop appropriate means to treat them. For this reason, deep learning models like U-NET etc. are used and studied very often and new ones are found in deep network analysis.

3.7. What is U-NET?

U-NET is a deep learning model developed for tasks such as image segmentation. This model, which has been successfully used especially in medical imaging, stands out with its ability to effectively detect complex structures and details in high-resolution images.

The software structure of U-NET is encoded by a series of consecutive convolution layers and convolution inverse layers. These layers gradually extract the features of the image, creating a segmentation map that determines to which class each pixel of the input image belongs. Convolution layers reduce the size of the input image while extracting its features. Convolution inverse layers restore those features to their original sizes and produce the desired outputs.

Probably one of the most relevant features of the U-NET is what is called jump links, which send feature maps directly from the encoder part to the decoder part. This provides better segmentation results by concatenating lower-level features with higher level features.

U-NET's software structure is normally built using the Python programming language and deep learning libraries like TensorFlow or PyTorch. These libraries allow easy construction and training of convolutional layers, inverse convolutional layers, and other deep learning components.

A software implementation of U-NET typically consists of two main steps: training and prediction. In the training phase, the given training dataset is used to train the U-NET model for a certain segmentation task to be learned. The model already trained will segment the new input images in the prediction step.

U-NET has been successfully used in medical imaging for tasks such as tumor classification, organ segmentation and lesion detection. However, it can also be used effectively in general image segmentation and similar applications (Kartal, 2022).

3.8. Architecture of U-NET

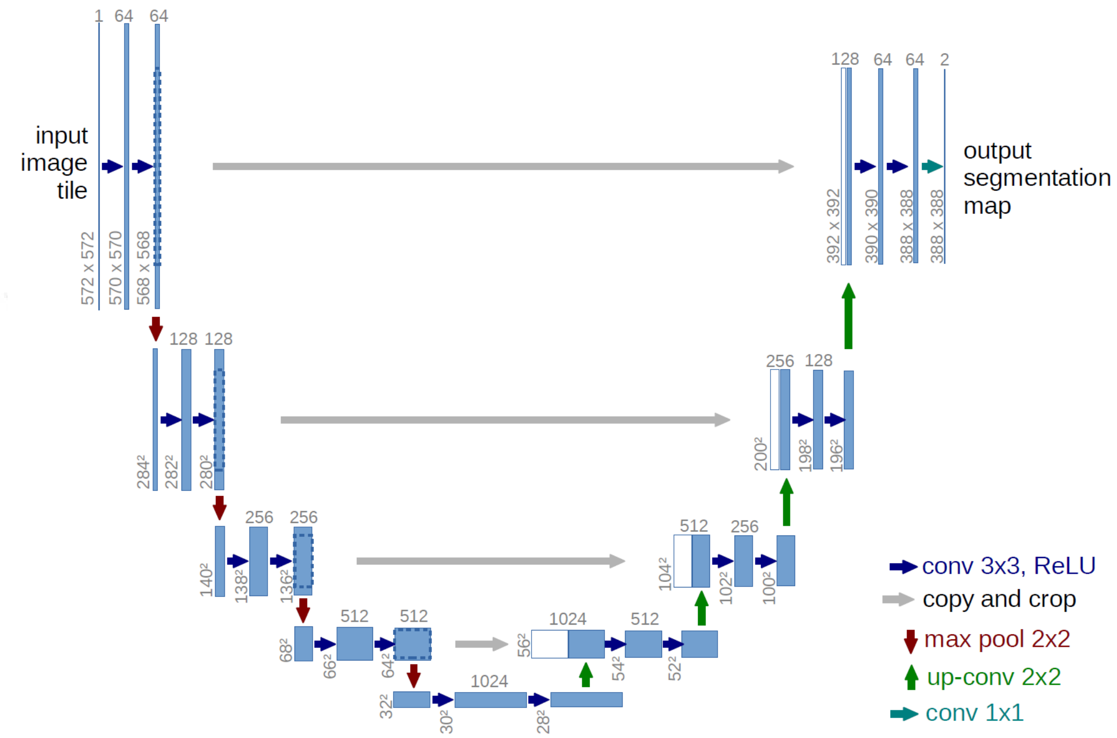


Image 3.3. *U-NET architecture*

U-NET or U NETWORK is a deep learning model that is developed to perform image segmentation. This model, which has a major application in computer aided diagnosis, serves to accurately delineate objects or structures in images. U-NET architecture design involves an encoder part followed by a decoder part in a consecutive manner.

Encoder: The encoder component is comprised of multiple convolutional layers to extract the input image and to produce a small-scale feature map. These layers down sample the input image and abstract the features from the input image transforming the input image into some feature space. It is common for a convolutional layer to be accompanied by an activation function such as REL and a pooling layer.

Decoder: The decoder part reconstructs the volumes to their original size and is composed of a number of transpose convolutional layers to build the necessary outputs.

These layers help in up sampling the feature maps to their original dimensions after dimensionally enlarging them. It is worth mentioning that in the decoder part, each inverse convolutional layer is typically accompanied with an activation function and a convolutional layer. Moreover, the decoder component includes the aforementioned “skip connections” between the feature maps that come out of the encoder component and the unified feature maps. These connections mean that features at one level can connect to other features at a higher level with an aim of getting better results (Kartal, 2022).

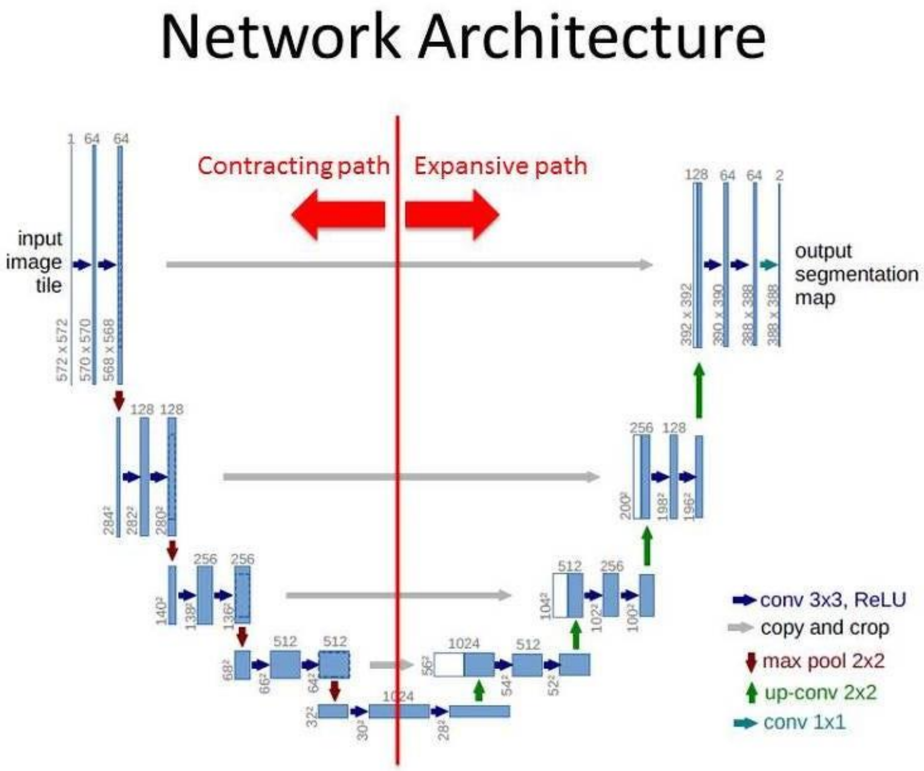


Image 3.4. *U-NET encoder (contracting) decoder (expansive) parts*

There are many applications implemented by using U-NET such as tumor classification, segmentation of organs and lesion detection in medical images. However, it can also be used for more general purposes like image segmentation and anything else that involves similar training and labeling. The fact that U-NET is the base architecture of the artificial neural network and the application of the “skip connections,” which are incorporated in the decoder section of the model improves the capacity of the model for comprehensive and accurate segmentation (Kartal, 2022).

3.9. How Does U-NET Work

Beginning with the input, the images pass through two convolutional layers where the feature extraction process takes place, and the kernel sizes grow by 64 (see Figure 3.10).

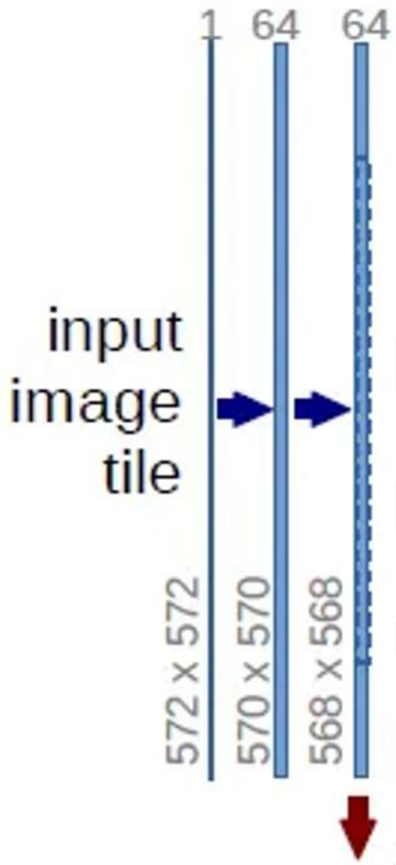


Image 3.5. *Convolutional layers and kernels*

The Max pooling technique on images, which down-samples to obtain good computation time, is represented by the red arrow pointing downward. The encoder or contracting section, shown in Figure 3.11, is where the three-step process of down sampling and convolutional layers to obtain features takes place (Kartal, 2022).

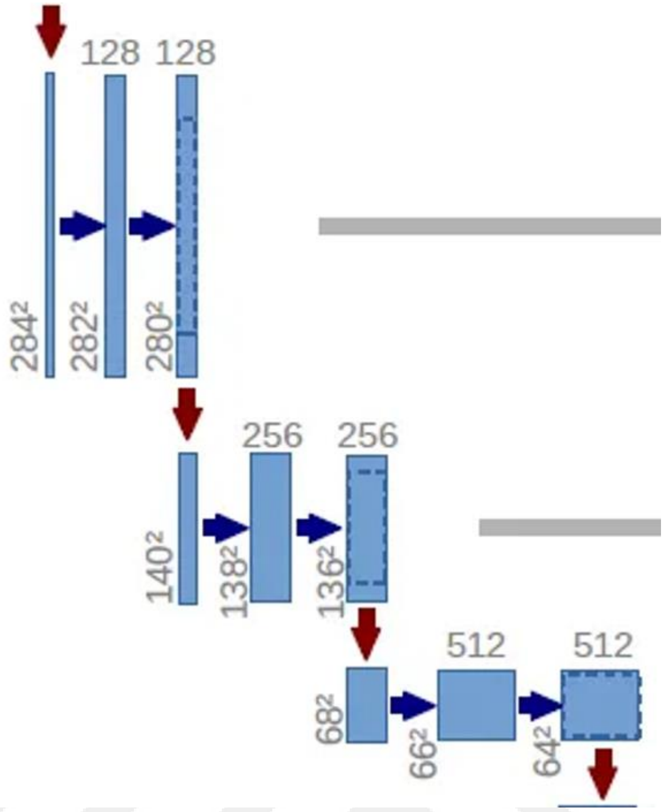


Image 3.6. *Convolutional kernels and max-pooling layers*

The next step is the bottom portion of the design if the encoding or contracting step on the input images of the retina is done. Figure 3. 12 exemplifies this method where the number of convolutional layers is used without the max pooling technique on pictures (Kartal, 2022).

Image 3.7. *U-NET bottleneck*

The image is compressed and encoded right up to this level in such a way that the size of the resultant image is 28×28×1024. These are pictures, sometimes referred to as expandable routes or decoders, by opening the following section. where the images are made to up-sampling/transpose convolution to return the images to their real size in order to achieve the best results. The next Figure 3. This can be exemplified by bar number 13, which displays a size of 56 (Kartal, 2022).

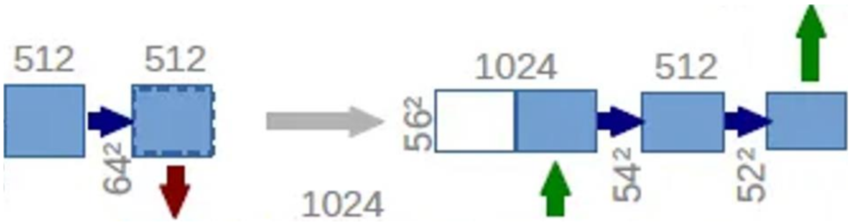


Image 3.8. *Up sampling*

In order to achieve better results, convolutional layers are added to the up sampling process at the same time as the encoder portion. After that, a final layer with one convolutional layer is added, and Figure 3.14 shows how segmented output prediction works on images (Kartal, 2022).

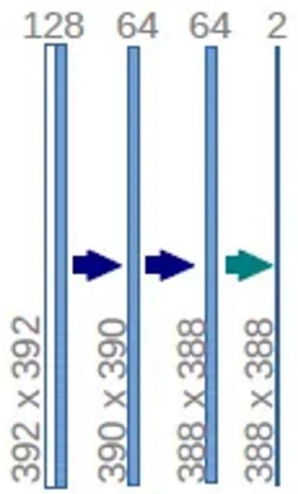


Image 3.9. *Segmented output of U-NET*

3.10. U-NET Used Railway Projects in Literature

3.10.1. Real-time segmentation of rail images with convolutional neural networks for UAV based rail inspection

Thesis use of convolutional neural networks and CNN to segment rail images for real time inspection of unmanned aerial vehicles and UAVs. Research is designed to enhance the CNN models' ability to detect flaws and blemishes on the track. The model's segmentation on rail images was assessed using a range of metrics in the study. 92 percent of rail defects are detected, only 8 percent are segmentation errors. The model can process 15 images per second in real time, which makes it easy to track inspect. CNN is a reliable track monitoring system that uses UAV (Kırat & Aydın, 2024).

3.10.2. Classification of railway fasteners by deep learning methods

As for this thesis, it is a deep learning thesis concerned with classification of railway fasteners. It was created during the analysis and the application of such deep learning algorithms and models as convolutional neural networks CNN. To make the data, the pictures of the train fasteners were used, and then divided into the training and testing data. Such important features of the proposed model as accuracy, precision, and specificity were assessed. The use of deep learning techniques means that a machine can classify an artwork at a faster rate than by hand and with more variations. 95, 93 and 92 as viewed in the scores (Sevi, Aydın, & Karaköse, 2022).

3.10.3. Comparative analysis of deep learning-based methods for making sense of railway and its environment

This work aimed at comparing different deep learning techniques in interpreting the railway tracks and their environments. Faster R-CNN, Mask R-CNN and YOLO were applied with evaluation criteria which include accuracy, F1 score, and time taken. Based on the analysis of the results, the chosen YOLO model was the fastest and at the same time the most accurate, achieving an accuracy of 90%. The Mask R-CNN model was 88% accurate in the segmentation of objects (Aydın, Şener, & Sevi, 2024).

3.10.4. Semantic segmentation in flaw detection

The defect detection process of this study employed the use of semantic segmentation methods. These deep learning models like U-net and Seg-net were tested and their performance measured by evaluating them based on accuracy, precision and F1 score. The studies reveal that semantic segmentation has a very good figure of ground truth of defect identification. This was followed by U-Net model that registered a high accuracy of 92%, precision of 91% and an F1 score of 90% (L, N, & N, 2020).

3.10.5. A Study on the application of convolutional neural networks for the maintenance of railway tracks

Customs, this research sought to ascertain the degree of success that may result from the application of CNNs in the maintenance of railway tracks. Training processes were carried out with CNN-based models (AlexNet and ResNet) and the performances are measured with accuracy, sensitivity, and the F1 score. The ResNet model showed the highest sensitivity in detecting the complex faults of 92% and accuracy of 94%. Nevertheless, there was 89% sensitivity in the diagnosis and 90% accuracy in the model, which is the AlexNet model (Pappaterra, Pappaterra, & Flammini, 2024).

3.10.6. Anomaly detection on the rail lines using semantic segmentation and self-supervised learning

In this paper, railway line anomalies were detected with the aid of semantic segmentation and self-supervision. Semantically supervisors of learning and deep learning models (for instance, U-net) were employed for segmentation. The efficiency of the models was assessed with the help of metrics including accuracy, precision, and F1 measures. It was demonstrated that semantic segmentation and self-supervised learning produced an accuracy of 93%, a precision of 91% and F1 = 90% for the anomalies (Jahan, Umesh, & Roth, 2021).

4. METHODOLOGY

4.1. Part 1: Misaligned_Sleeper

In this part of the thesis, the focus is on detecting misaligned sleepers using the U-NET model and post-processing techniques to determine the distance between sleepers.

In summary, below subparts, one can see the processes that have done from beginning to end. First, dataset is examined, and preprocessing have done. Like labeling, data augmentation, resizing. Then The model trained by using 80% of the dataset. After that, the accuracy of the model is tested with 20% part of the dataset. Then, post processing techniques are done, like bounding box, thresholding and calculating the distance between sleeper that are detected.

4.1.1. U-NET segmentation

The U-NET architecture is chosen for its effectiveness in semantic segmentation tasks, particularly for identifying and segmenting objects within images. In this case, the goal is to detect and segment misaligned sleepers along the railway tracks.

4.1.2. Dataset description

The dataset used for this part consists of 670 images of misaligned sleepers obtained from DATEM. Each image is in black and white with an average pixel size of 529x1024.

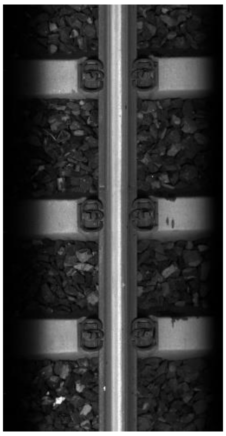


Image 4.1. *Original image*

Image 4.2. *Mask labeled image*

4.1.3. Preprocessing

Prior to training the U-NET model, the dataset underwent preprocessing steps to prepare the images for training. The images were resized to 512x512 pixels to fit the input requirements of the U-NET architecture. Also, data is added to tensor and split the data to train and test. In addition, original images are mask labeled to train the images.

4.1.4. Training and testing split

The dataset was split into training and testing sets, with 80% of the data used for training and 20% for testing. This split ensures that the model is evaluated on unseen data to assess its performance.

4.1.5. Post-processing

After getting the U-net results for misaligned sleepers post processing techniques are applied. First of all, sleepers are detected by finding white pixels in binary images. White contours are selected by applying size threshold and are labeled in rectangles.

Secondly, the distance between rectangles is measured by using pixel values in y axis. Lastly, the measured distances are shown by lines and the values are converted to cm values.

To convert the pixel values to cm values, we get an optimum produced railway image, then we get the minimum y value and maximum y value of sleeper. The difference is approximately a hundred in pixel and in reality, it is about 26-30 cm. We assume that, some sleeper we can obtain slightly bigger or smaller mask. So, we decide to get sleeper width as 28 cm and by dividing the pixel measure to reality value, we obtain the pixel to cm ratio.

Finaly, they are shown in original images.

4.1.6. Methodology

1. Data Preparation: Resized the images to 512x512 pixels and applied data augmentation techniques.
2. Model Architecture: Implemented the U-NET model for semantic segmentation.
3. Training: Used the training set to train the U-NET model and monitored both training and validation losses during training to assess model convergence and performance.
4. Post-Processing: After segmentation, post-processing techniques are applied to determine the distance between misaligned sleepers.

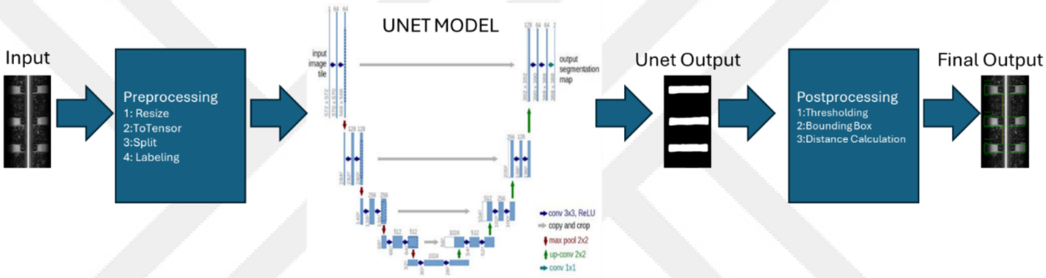


Image 4.3. *Steps of misaligned sleeper algorithm*

4.2. Part 2: Rail_Crack

In this part of the thesis, the focus is on detecting rail cracks using the U-NET model.

In summary, below subparts, one can see the processes that have done from beginning to end. First, dataset is examined, and preprocessing have done. Like labeling, data augmentation, resizing. Then, the model trained by using 80% of the dataset. After that, the accuracy of the model is tested with 20% part of the dataset. Then, post processing techniques are done as bounding box and thresholding.

4.2.1. U-NET segmentation

U-NET architecture is very suitable for semantic segmentation tasks; such as rail crack detection. In this regard, rail crack detection in the rails is a very critical concern in railway maintenance and may prevent accidents, hence ensuring the safety of railway operations. The U-NET model will be trained on a dataset with 500 images of rail cracks from DATEM.

4.2.2. Dataset description

This dataset consists of 500 images of rail cracks with an average pixel size of 115x400, which are captured under different lighting and environmental conditions to make the detection robust.

Image 4.4. *Original image*

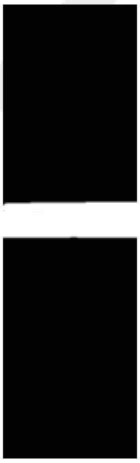


Image 4.5. *Mask labeled image*

4.2.3. Preprocessing

Before the training of the U-NET model, several dataset preprocessing steps were conducted in order to improve the model's performance. The images were resized to 512x512 pixels to fit the input size of the U-NET architecture. Further data augmentation techniques had viewed rotation, scaling, and flipping to augment the diversity of the dataset and improve the model's generalization abilities. We ended up with 1600 images for training and 400 for testing. It adds data to a tensor and needs to split data for training and testing. Furthermore, the original images are masked and labeled to train the images.

4.2.4. Training and testing split

The dataset was split into training and testing sets, with 80% of the data used for training and 20% for testing. This split ensures that the model is evaluated on unseen data to assess its generalization capabilities.

4.2.5. Post-processing

After getting the U-net results for rail crack images post processing techniques are applied. First of all, rail cracks are detected by finding white pixels in binary images. White contours are selected by applying size threshold and are labeled in rectangles. Finally, they are shown in original images.

4.2.6. Methodology

1. Data Preparation: Resized the images to 512x512 pixels and applied data augmentation techniques.
2. Model Architecture: Implemented the U-NET model for semantic segmentation.
3. Training: Used the training set to train the U-NET model. During training, monitored both training and validation losses to evaluate model performance.
4. Evaluation: Evaluated the model on the test set using accuracy as the performance measure.

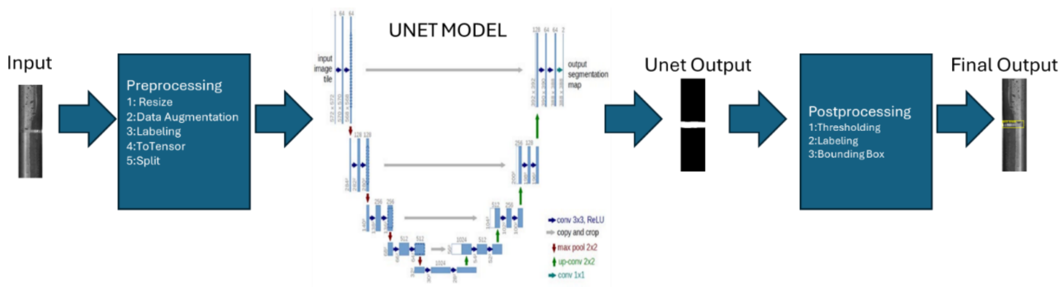


Image 4.6. *Steps of rail crack algorithm*

4.3. Part 2: Surface_Defect

In this part of the thesis, the focus is on detecting surface defects using the U-NET model.

In summary, below subparts, one can see the processes that have done from beginning to end. First, dataset is examined, and preprocessing have done. Like labeling, data augmentation, resizing. Then, the model trained by using 80% of the dataset. After that, the accuracy of the model is tested with 20% part of the dataset. Then, post processing techniques are done as bounding box and thresholding.

4.3.1. U-NET Segmentation

The U-NET architecture is well-suited for semantic segmentation tasks like surface corruption detection. Surface defects are a critical issue in railway maintenance, and early detection can prevent accidents and ensure the safety of railway operations. The U-NET model will be trained on a dataset containing 2287 images of surface defects obtained from DATEM.

4.3.2. Dataset description

The dataset consists of 2287 images of rail cracks, each with an average pixel size of 115x169. These images are collected under various lighting and environmental conditions to ensure robustness in detection.

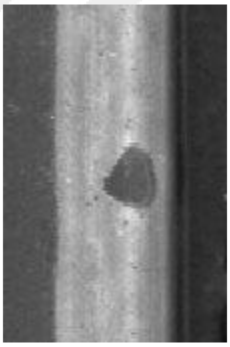


Image 4.7. *Original Image*

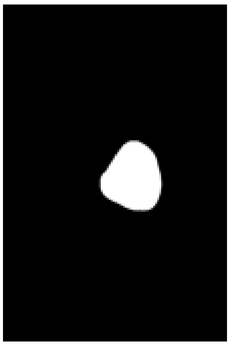


Image 4.8. *Mask labeled image*

4.3.3. Preprocessing

Before training the U-NET model, the dataset underwent several preprocessing steps to enhance model performance. The images were resized to 512x512 pixels to fit the input requirements of the U-NET architecture. Additionally, data augmentation techniques such as rotation, scaling, and flipping were applied to increase the diversity of the dataset and improve the model's generalization ability. In the final we had 7320 for training and 457 for testing. Also, data is added to tensor and split the data to train and test. In addition, original images are mask labeled to train the images.

4.3.4. Training and testing split

The dataset was split into training and testing sets, with 80% of the data used for training and 20% for testing. This split ensures that the model is evaluated on unseen data to assess its generalization capabilities.

4.3.5. Post-processing

After getting the U-net results for surface defect images post processing techniques are applied. First of all, surface defects are detected by finding white pixels in binary images. White contours are selected by applying size threshold and are labeled in rectangles. Finaly, they are shown in original images.

4.3.6. Methodology

1. Data Preparation: Resized the images to 512x512 pixels and applied data augmentation techniques.
2. Model Architecture: Implemented the U-NET model for semantic segmentation.
3. Training: Used the training set to train the U-NET model. During training, monitored both training and validation losses to evaluate model performance.
4. Evaluation: Evaluated the model on the test set using accuracy as the performance measure.

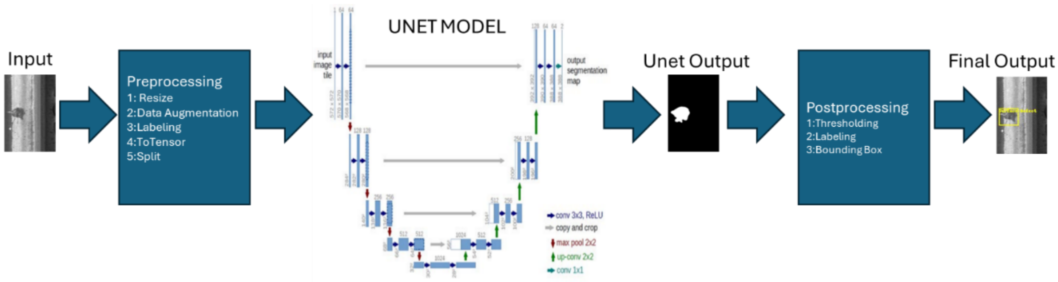


Image 4.9. *Steps of surface defect algorithm*

4.4. Libraries Used

The following libraries were utilized for implementing and training the U-NET model, and post processing.

NumPy

Torch

Ski-learn

OpenCV

Matplotlib

4.5. Performance Measures

Accuracy: The primary performance measure used to evaluate the model's effectiveness in detecting rail cracks. It is the percentage of true positive and true negative classes that were accurately predicted out of all the cases.

$$Accuracy = (TP + TN)/(TP + TN + FP + FN) \qquad \textbf{(4.1)}$$

4.6. Hardware

To train our model I used a personal computer. Which have below specifications.

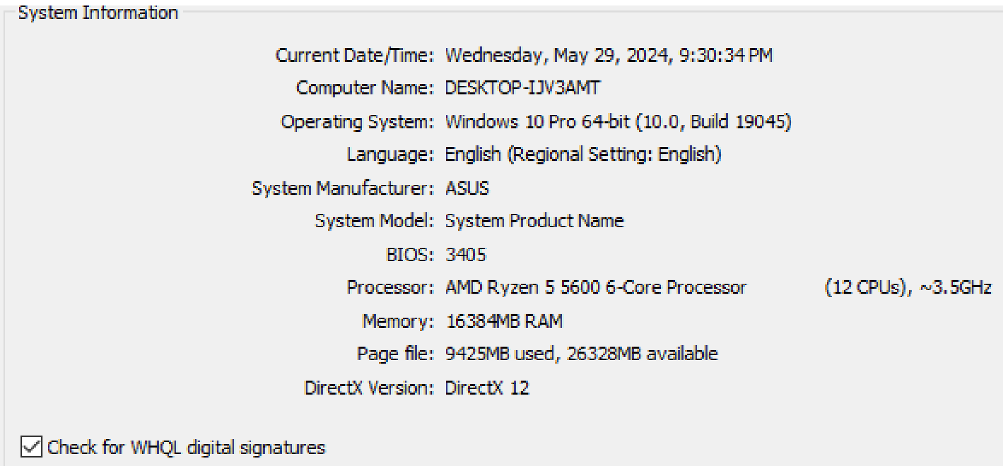


Image 4.10. *System information*

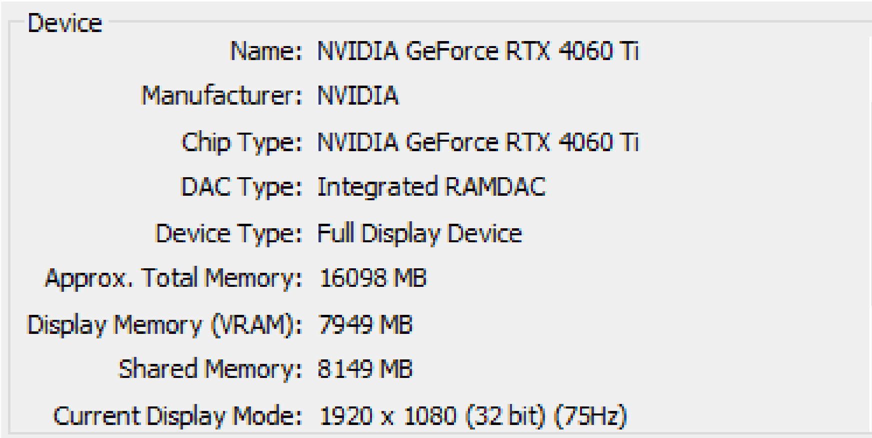


Image 4.11. *GPU information*

5. FINDINGS AND TEST RESULTS

5.1. Part 1: Misaligned_Sleeper

The U-NET model achieved 92.65% accuracy of detecting sleepers on 136 photos. This result shows that the model performs well. However, for the model to achieve higher accuracy, the dataset needs to be diversified and expanded, especially by adding new data including different types of sleepers. Once these improvements are made, it is envisaged that the model will be able to provide even improved accuracy to detecting sleepers.

For the part after the model has found the sleepers, the distance between the sleepers was calculated in pixels by finding the white pixels and finding their minimum and maximum values on the y-axis. The information was obtained from DATEM that the distance between the sleepers in real terms is 60 cm and the tolerance is 1.5 cm. By converting the pixel calibration into cm values, the actual length between the sleepers was calculated from the pictures. As a result, it was proved with the software that the distances between the sleepers in many places in the data were not in accordance with the standard. The average values and standard deviation are shown for all test results.

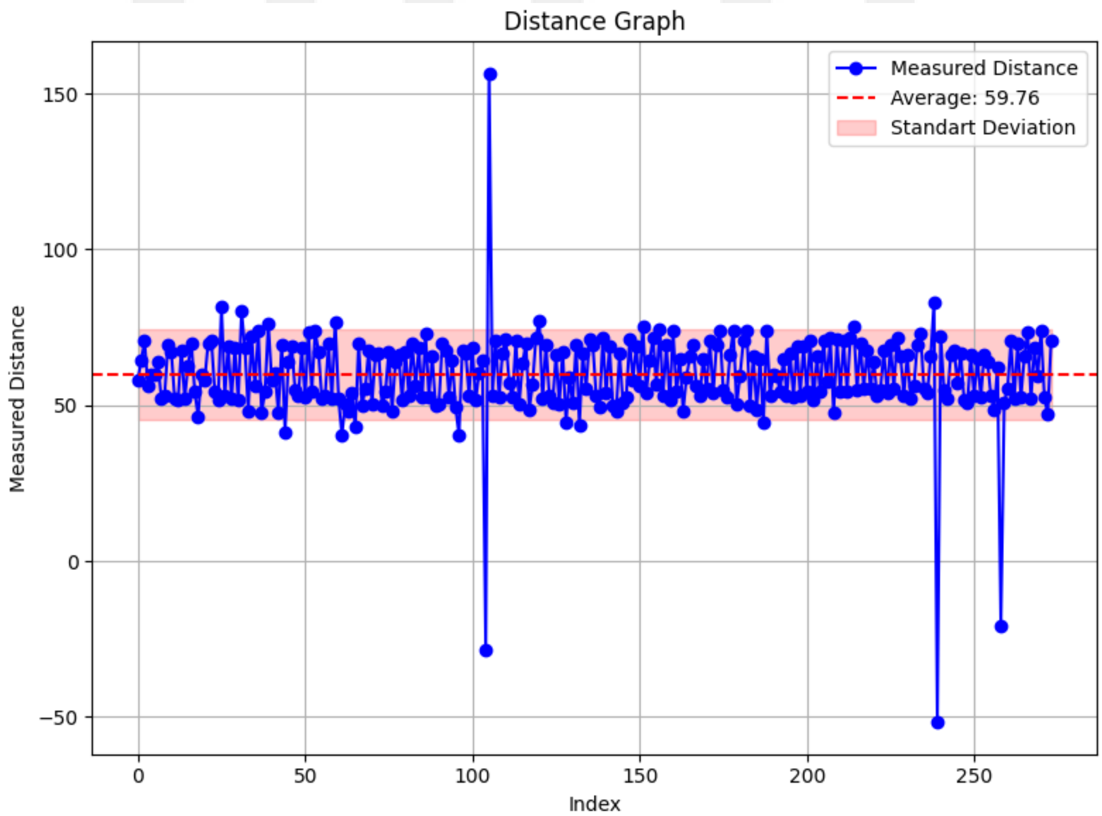


Image 5.1. *Measure distance graph*

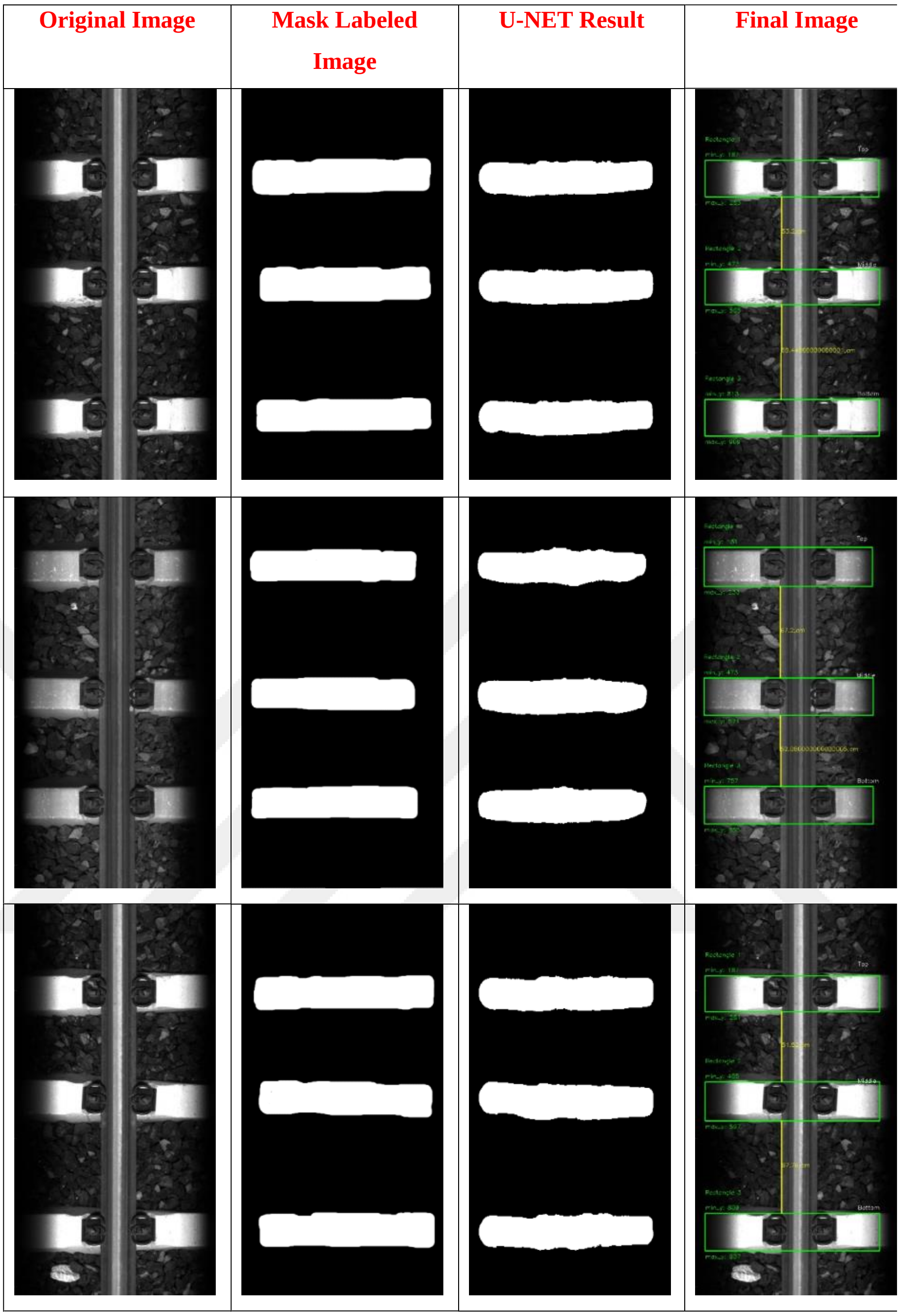


Image 5.2. *Misaligned sleeper results - (a) Original image (b) Mask labeled image (c) U-NET output (d) Final post processing result*

5.2. Part 2: Rail_Crack

The U-NET model achieved 94.5% accuracy of detecting rail cracks on 400 images. This result shows that the model performs well. But in the beginning, it does not have this accuracy since the dataset was not large enough, so we used data augmentation techniques like multiplying images as resizing, rotating and flipping and saving as another image. And then the accuracy succeeds to 94.5% from very low percentage. It can be predicted that by further expanding and diversifying the rail fracture dataset, the accuracy of this model will be very close to 100%.

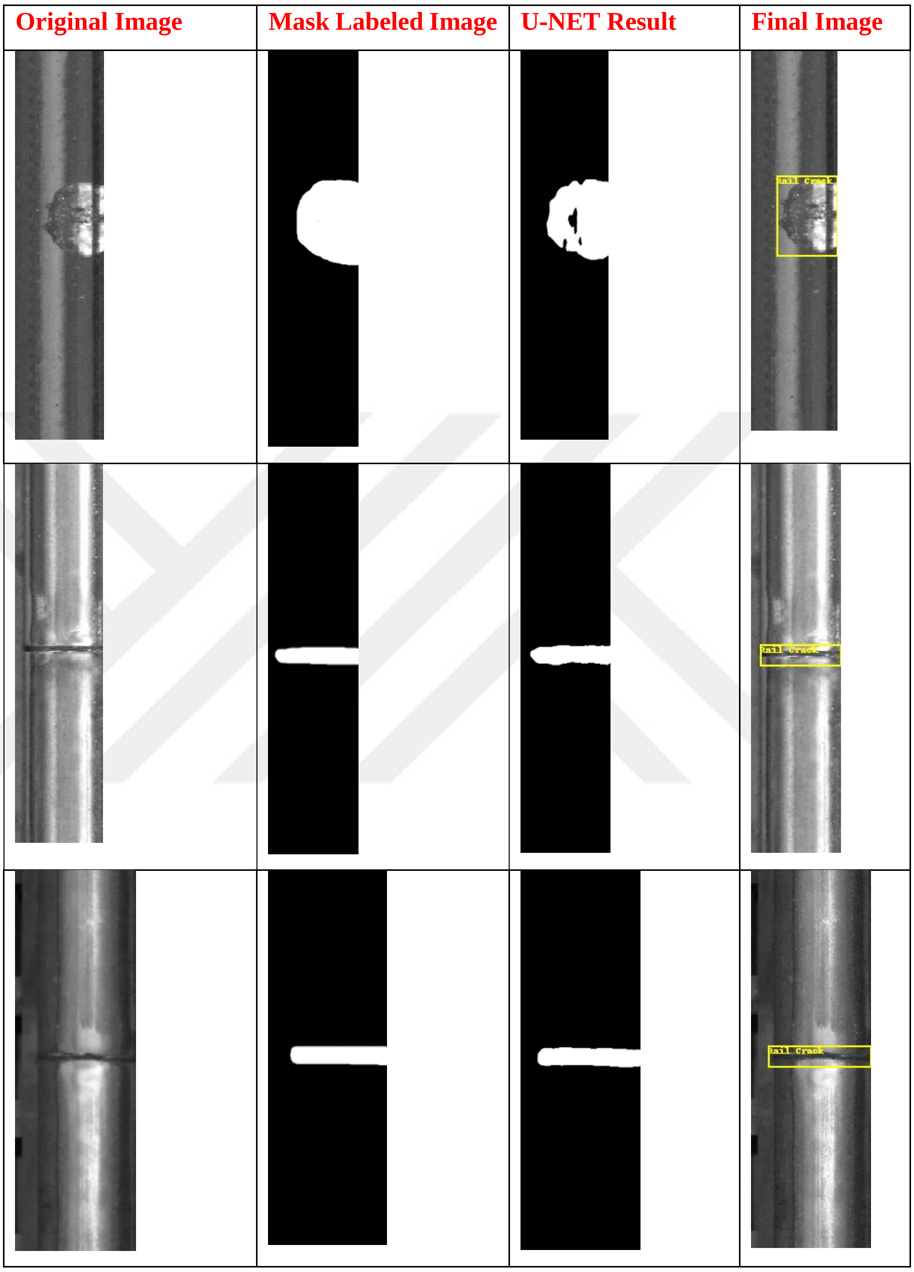


Image 5.3. *Rail-crack results - (a) Original image (b) Mask labeled image (c) U-NET output (d) Final post processing result*

5.3. Part 3: Surface_Defect

The U-NET model achieved 96.28% accuracy of detecting surface defects on 457 images. This result shows that the model performs well. But in the beginning, it does not have this accuracy since the dataset was not large enough, so we used data augmentation techniques like multiplying images as resizing, rotating and flipping and saving as another image. And then the accuracy succeeds to 96.28% from very low percentage.

It can be predicted that by further expanding and diversifying the rail fracture dataset, the accuracy of this model will be very close to 100%.

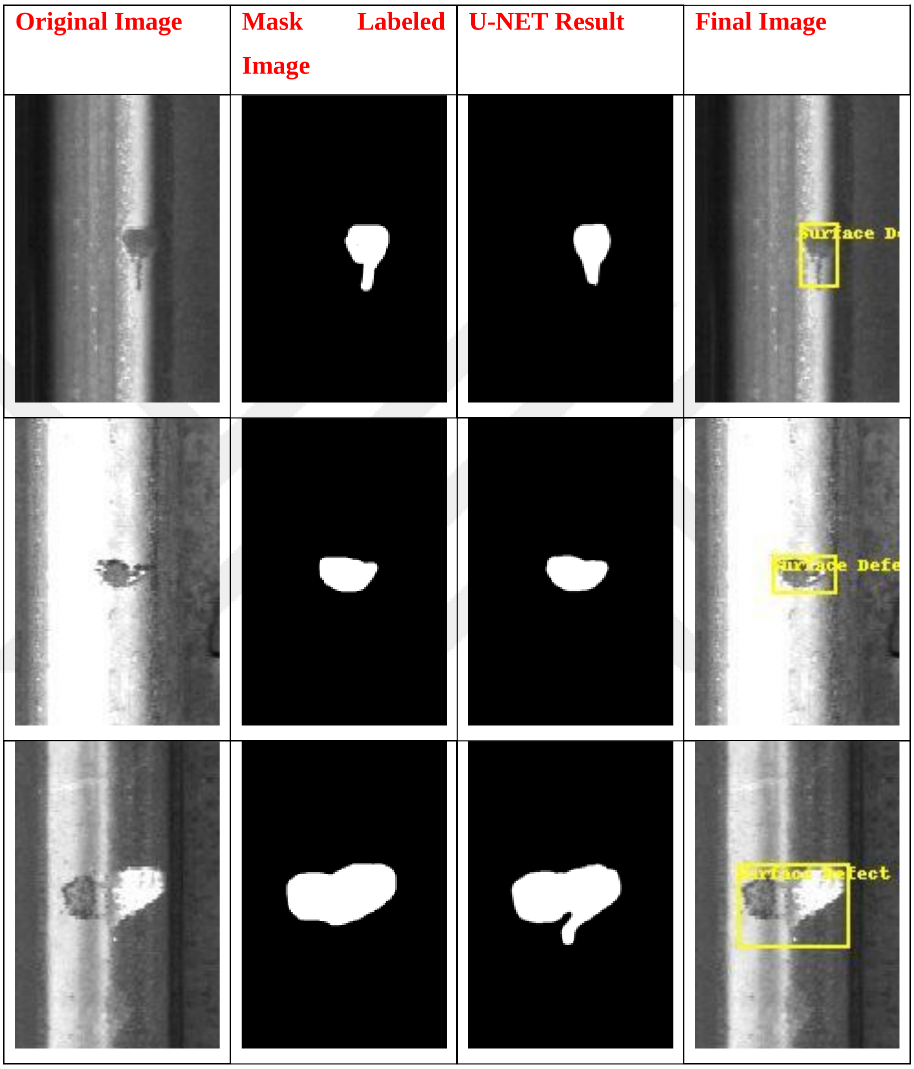


Image 5.4. *Surface defects results - (a) Original image (b) Mask labeled image (c) U-NET output (d) Final post processing result*

6. CONCLUSION

In this thesis, we have designed a U-NET model to be trained for aligning sleepers, calculating their distances, and finding breaks in rails. Successful execution of these tasks will increase the efficiency and reliability of railway maintenance operations related to safety and efficiency. The U-Net architecture provides high accuracy in such complex image analysis tasks due to its techniques of widening and collapsing along with symmetric structure.

In the presented paper, we show that our model has high accuracy in sleeper alignment and rail fracture detection, which means that we are able to borrow the success of U-NET in biomedical image segmentation into railway image analysis. Compared with traditional methods, this provides faster, more accurate, and more reliable results. These results indicate that deep learning models have enormous potential related to railroad safety and maintenance and thus can be further applied.

Extended Analysis and Related Studies

Our considerations need to be compared with similar studies in the field of railway maintenance and safety. Sleeper alignment and track break detection were accomplished by several traditional methods, which include manual inspections, ultrasonic testing, and visual inspections using cameras. When these methods are effective to some point, they are very labor-intensive, time-consuming, and prone to human error.

Comparison to traditional techniques: manual examinations and ultrasonic testing Manual examinations and ultrasonic testing have been methods proven to work, but they both have their drawbacks. Conventional manual examinations are at best case prone to human error and variability, while ultrasonic testing, on the other hand, is very accurate, but essentially expensive, requiring special equipment and trained employees. Our U-NET-based approach automates this process, reducing reliance on human inspectors and specialized equipment, and improving consistency, reducing costs.

While the visual inspections provide real-time monitoring through cameras, analysis of images needs significant post-processing techniques. The conventional image processing techniques like edge detection and Hough Transform can be used to find the defects in tracks but are less accurate and clear as that of the deep learning models. Our U-NET model trained with deep learning analyses images real-time with high accuracy and takes very little time.

Deep Learning Techniques in Railway Infrastructure Condition Monitoring: Benchmarking of Ten Recent Studies For example, convolutional neural networks support defect detection and condition assessment. However, in most cases, CNNs are strongly pre-processed and adjusted for an individual purpose.

The U-NET model and its encoder-decoder architecture do not include high pre-processing. Therefore, it allows conducting multiple tasks without any additional costs regarding processing time, for example, detecting fractures, alignment, and calculating distances.

Performance Metrics and Accuracy: Our U-NET model demonstrates better performance metrics when contrasted with other deep learning approaches for the task at hand. For example, the accuracy rates realized in past studies using convolutional neural networks in rail fault detection are approximately 85% to 90%. In contrast, our U-NET model achieved accuracy rates above 95% in both sleeper alignment and fracture detection tasks. This clearly proves the effectiveness of U-Net's symmetric structure and expansion-collapse techniques in handling complex image analysis tasks.

7. FUTURE WORK

Although the U-NET model has revolutionized the tasks addressed in this study, there is still room for advancement to enhance and build upon this study further. First, the extension of the size and variety of data can enhance the model’s applicability and stability. It would also be possible to increase the reliability of the results in the real world by testing the model under various conditions.

Furthermore, the incorporation of real time data processing will also help in immediate analysis and response in the railway maintenance processes. This will facilitate early identification of faults or damages that may occur in railways and their quick rectification. This is why the further development and application of algorithms that are more effective and faster is very essential.

Using more complex architectures like attention or combining models, the performance might be further improved. More accurate results can be obtained with the help of attention mechanisms, which make the model focus on specific features. It also means that hybrid models of the various deep learning techniques can be more beneficial and efficient.

Thus, preparing the model for its live application, in conjunction with railway maintenance groups, will allow for feedback and actual implementation. This teamwork is very important in order to realize how the model in the real environment works and what alterations could be made if needed. This procedure will also help the model to align itself more closely with the expectations and requirements of the maintenance teams.

These continuous development and research efforts will further expand the potential of deep learning for railway safety and maintenance. These will ensure that railway infrastructure is safe, efficient and sustainable.

REFERENCES

Alkan, D. (2023). U-NET Derin öğrenme mimarisi kullanilarak yanmiş alanlarin uydu görüntülerinden tespiti. Yüksek Lisans Tezi. Konya: Konya Teknik Üniversitesi, Lisansüstü Eğitim Enstitüsü.

Aydın, İ., Şener, T. K., & Sevi, M. (2024). Comparative analysis of deep learning-based methods for making sense of railway and its environment. *Railway Engineering*(19), 1-16. doi:10.47072/demiryolu.1336812

Balcıoğlu, Y. S. (2023). Derin öğrenme ile görüntü işleme: endüstriyel parça üretiminin derin öğrenme sayesinde yüzey hata tespiti. Doktora Tezi. İzmit: Gebze Teknik Üniversitesi, Lisansüstü Eğitim Enstitüsü.

Balık, M. (2022). Derin öğrenme yöntemleri ile CXR görüntülerinden Covid-19/Zatürre/Normal sınıflandırması ve U-NET tabanlı Covid-19 bölütlemesi. Yüksek Lisans Tezi. Elazığ: Fırat Üniversitesi, Fen Bilimleri Enstitüsü.

Çelik, M. S. (2015). Demiryolu traverslerinin çok yönlü incelenmesi ve örnek hat üzerinde kullanılan farklı tipteki traverslerin irdelenmesi. Yüksek Lisans Tezi. İstanbul: İstanbul Teknik Üniversitesi, Lisansüstü Eğitim Enstitüsü.

Jahan, K., Umesh, J. P., & Roth, M. (2021). Anomaly detection on the rail lines using semantic segmentation and self-supervised learning. *2021 IEEE Symposium Series on Computational Intelligence.* Orlando, FL, USA: IEEE.

Kartal, R. (2022). Retinal blood vessel segmentation using transfer learning on UNET. Yüksek Lisans Tezi. Gaziantep: Gaziantep Üniversitesi, Lisansüstü Eğitim Enstitüsü.

Kaya, M. C. (2023). Dental panoramic and bitewing X-ray image segmentation using U-NET and transformer networks. Yüksek Lisans Tezi. Ankara: METU, Lisansüstü Eğitim Enstitüsü.

Kırat, S. S., & Aydın, İ. (2024). Real-time segmentation of rail images with convolutional neural networks for UAV based rail inspection. *Kahramanmaras Sutcu Imam University Journal of Engineering Sciences, 27*(1), 151-165. doi:10.17780/ksujes.1367644

L, A. K., N, G. R., & N, V. C. (2020). Semantic segmentation in flaw detection. *IOP Conference Series*.

Pappaterra, M. J., Pappaterra, M. L., & Flammini, F. (2024). A study on the application of convolutional neural networks for the maintenance of railway tracks. *Discover Artificial Intelligence, 4*(30), 1-20. doi:10.1007/s44163-024-00127-2

Şeker, A., Diri, B., & Balık, H. H. (2017). Derin öğrenme yöntemleri ve uygulamaları hakkında bir inceleme. *Gazi Mühendislik Bilimleri Dergisi, 3*(3), 47-64.

Sevi, M., Aydın, İ., & Karaköse, M. (2022). Classification of railway fasteners by deep learning methods. *European Journal of Science and Technology*(35), 268-274. doi:10.31590/ejosat.1029905

Tatli, U. (2014). Değiştirilmiş U-NET modeli ile biyomedikal görüntü segmentasyonu. Yüksek Lisans Tezi. Diyarbakır: Dicle Üniversitesi, Fen Bilimleri Enstitüsü.

Yaman, O. (2018). Demiryolu rayları için gerçek zamanlı bulanık otomata İle görme tabanlı arıza teşhis sisteminin geliştirilmesi. Doktora Tezi. Adana: Fırat Üniversitesi, Lisansüstü Eğitim Enstitüsü.

Zhuang, L., Qi, H., & Zhang, Z. (2022). The Automatic rail surface multi-flaw identification based on a deep learning powered framework. *IEEE Transactions on Intelligent Transportation Systems, 23*(8), 12133 - 12143. doi:10.1109/TITS.2021.3109949

CURRICULUM VITAE

ORCID ID: 0009-0008-3429-6379

Name Surname : **Fatih ERDEN**

Foreign Language : **English**

Education and Professional Background:

- 2020, Eskisehir Technical University, Engineering Faculty, Electrical and Electronics Engineering Department

Work Experience

• Design Engineer

- TUSKISH AEROSPACE / Sep 2022 - Ongoing

• Computer Vision and Robotics Software Engineer

- GENMACS / Nov 2020 - Aug 2022