**POWERLINE DETECTION FOR
AIRCRAFT FLIGHT SAFETY
USING IMAGE PROCESSING**

Ph. D. Dissertation

Ömer Emre YETGİN

Eskişehir, 2018

POWERLINE DETECTION FOR AIRCRAFT FLIGHT SAFETY
USING IMAGE PROCESSING

Ömer Emre YETGİN

Ph.D. DISSERTATION

Department of Electrical and Electronics Engineering

Supervisor : Prof. Dr. Ömer Nezih GEREK

Eskişehir

Anadolu University

Graduate School of Sciences

November, 2018

FINAL APPROVAL FOR THESIS

This thesis titled "**POWERLINE DETECTION FOR AIRCRAFT FLIGHT SAFETY USING IMAGE PROCESSING**" has been prepared and submitted by Ömer Emre YETGİN in partial fulfillment of the requirements in "Anadolu University Directive on Graduate Education and Examination" for the Degree of Doctor of Philosophy (Ph.D.) in Electrical and Electronics Department has been examined and approved on 23/11/2018.

Committee Members		**Signature**
Member (Supervisor) :	Prof. Dr. Ömer Nezih GEREK	....
Member :	Prof. Dr. Rıfat EDİZKAN	....
Member :	Assoc. Prof. Dr. Kemal ÖZKAN	....
Member :	Assoc. Prof. Dr. Tansu FİLİK	....
Member :	Assist. Prof. Dr. Cihan TOPAL	....

Prof. Dr. Ersin Yücel

Director of Graduate School of Sciences

ABSTRACT

POWERLINE DETECTION FOR AIRCRAFT FLIGHT SAFETY
USING IMAGE PROCESSING

Ömer Emre YETGİN

Department of Electrical and Electronics Engineering
Anadolu University, **Graduate School of Sciences, November, 2018**

Supervisor: Prof. Dr. Ömer Nezih GEREK

In this thesis, several active methods have been proposed for the safety of air vehicles, which prevent air vehicles from colliding with electrical wires. In this context, Discrete Cosine Transform, Linear Binary Pattern, Histogram of Gradient and Convolutional Neural Network methods were used. The methods were tested on image databases, which consist of real life aerial images that were captured and curated by ourselves. The recognition performances of such scenes with these methods were compared. The effects of various pre-processing / pre-training on the classification and feature extraction steps were also examined.

Keywords: Powerlines, Discrete cosine transform, Linear binary pattern, Histogram of gradients, Convolutional neural network.

ÖZET

HAVA ARACI UÇUŞ GÜVENLİĞİ İÇİN GÖRÜNTÜ İŞLEME KULLANILARAK GERİLİM TELİ TESPİTİ

Ömer Emre YETGİN

Elektrik-Elektronik Mühendisliği Anabilim Dalı
Anadolu Üniversitesi, **Fen Bilimleri Enstitüsü, Kasım, 2018**

Danışman: Prof. Dr. Ömer Nezih GEREK

Bu tezde, hava araçlarının güvenliğine yönelik olarak, hava araçlarının elektrik tellerine çarpmalarını engeleyecek aktif birkaç yöntem önerilmiştir. Bu kapsamda, Ayrık Kosinüs Dönüşümü, Lineer İkili Desen, Gradyenlerin Histogramı ve Evrişimsel Nöral Ağ yöntemleri kullanılmıştır. Kullanılan yöntemler, tarafımızdan oluşturulan ve gerçek görüntülerden oluşturulan veritabanı üzerinde denenmiştir. Bu yöntemlerin elektrik teli içeren sahneleri tanıma performansları karşılaştırılmıştır. Yöntemler uygulanmadan önce, ön-işleme / ön-eğitim işlemlerinde hangi metodların kullanılacağı ve bu metodların performansları da incelenmiştir.

Anahtar Kelimeler: Yüksek gerilim telleri, Ayrık kosinüs dönüşümü, Lineer ikili desen, Gradyenlerin histogramı, Evrişimsel nöral ağ.

23/11/2018

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 Eskisehir 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.

Ömer Emre YETGİN

ACKNOWLEDGEMENTS

First and foremost I thank my supervisor Prof. Dr. Ömer Nezih GEREK for providing to study and complete my Ph.D. dissertation. I am extremely grateful for our many discussions which helped guide me in the right direction, his motivating attitude, his availability as a supervisor and his many insightful suggestions.

And most importantly I would like to thank you my family for their encouragement and support and for keeping me honest.

<div align="right">

Ömer Emre YETGİN

November, 2018

</div>

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ACRONYMS

CAA	: Civil Aviation Authority
CNN	: Convolutional Neural Network
CS	: Classical Selection
DCT	: Discrete Cosine Transform
FFT	: Fast Fourier Transform
FP	: False Positive
GPS	: Global Positioning System
GPU	: Graphics Processing Unit
HELLAS	: Helicopter Laser Radar System
HOG	: Histogram of Gradient
IR	: Infrared
ILSVRC	: ImageNet Large Scale Visual Recognition Challenge
JPEG	: Joint Photographic Experts Group
LBP	: Linear Binary Pattern
LOAS	: Laser Obstacle Awareness System
LOG-DCT	: Logarithm of DCT
OASYS	: Obstacle Avoidance Laser Radar System
PB	: Patch-Based
ROC	: Receiver Operating Characteristic
RS	: Reversed Selection
SVM	: Support Vector Machines
TEIAS	: Turkey Electricity Transmission Company
TP	: True Positive
VL	: Visible Light
WSPS	: Wire Strike Protection System
2-D	: 2 Dimension

1. INTRODUCTION

Flight safety for air vehicles is one of the most important issues. There are many factors affecting flight safety. One of the most important of these factors are electrical wires. For this reason, it is of great importance to develop a system that will warn the pilots beforehand by detecting which may pose great danger for certain aircrafts, such as helicopters.

Wire accidents usually occur during low altitude aircraft flights. Helicopters and small aircraft pilots have difficulty in detecting thin electrical wires when flying low. A low-flying aircraft is at danger because it is close to obstacles such as trees, electrical wires, buildings and radio poles. Furthermore, in urban areas, objects with linear patterns such as buildings, trees and shadows on helicopters' flight routes make the detection of electrical wires even more difficult.

Due to the low contrast between the cable and the background, especially in cloudy weather, the detection of thin objects such as electrical wires become a very difficult problem and several studies have been carried out on this problem (Yetgin, 2013, Liu, 2014, Song, 2014).

Conditions where the wire cannot be distinguished from the background are a major danger to the aircraft. Often, the only electrical wires in the most unexpected places in the countryside are the biggest danger and it is very difficult to detect these wires from the aircraft.

Factors affecting the visibility of the wires are the position of the sun, the background color of the wire, the terrain obstacles and the bad weather conditions. In addition, the correct detection of the wire can be difficult due to various optical events. When the low-lying wires are displayed together with the high-pitched wires, they appear to be even farther away. This effect is improved when these wires are approached closer.

Another factor that prevents the pilot's ability to detect electric poles is the physiological limitations of the eye. When looking forward, the normal viewing angle of each eye is approximately 120 degrees in the vertical and 200 degrees in the horizontal. However, the field of view, which provides a clear detail of the electrical wires, is rather narrow. Therefore, the detection of electrical wires by the human eye becomes a very difficult problem.

In addition to the physical conditions mentioned above, there are many human factors such as stress, physical and mental fatigue. However, the most important human factor associated with low flight missions is the distraction of the pilot. There are several reasons for the distraction of the pilot. These are bad weather, personal stress, objects on the ground, radio calls, disturbances in hardware and passengers.

Distractibility generally be divided into four classes:
1. Visual distributions-especially eye-catching images
2. Audio distributions - radio or telephone
3. Physical distributions-disease
4. Mental distributions-diving into the tHought

These distractions can take place one by one or together and turn the pilot's attention away from the flight.

The pilot's view of the outside environment is also affected by the shadow and vibration in the cockpit. Shadows often reduce the quality of vision. As is known, human performance decreases after exposure to a certain level of vibration for a long time. For example, the natural frequency of the human eye and internal structure ranges from 20 to 90 Hz. Similarly, the structural frequency of many helicopters is in the range of 20 Hz.

Although power poles are one of the most reliable indicators for wire detection, in some cases it is not easy to see these poles. In particular, wooden poles can be camouflaged with landscapes or greenery or trees. For example, since the color of oxidized copper wires turns green, it is very difficult to separate these wires from the grasses and trees in the background.

Various exciter objects can be placed near the wire to indicate the presence of electrical wires. In order to warn the pilots in the current situation, 2 colored (white and orange) and 4 kg weighted aluminum spheres are placed on high voltage cables (Yetgin, 2017).
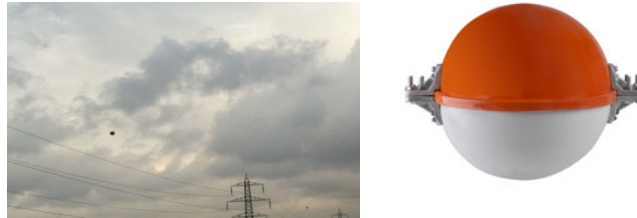
Figure. 1.1. *Alert sphere is an example app and warning ball (Yetgin, 2017).*

Pilots are given various trainings on avoidance of wire accidents, and helicopter pilots can avoid wire accidents with various maneuvers. Experts thinks that the area to be operated before the low flight operations will be explored from a higher flight. However, this is not always possible.

1.1. Alert Response Time Measurement

Pilots can save a helicopter as soon as possible and in an emergency. For this purpose, it is necessary to determine the operational characteristics of the pilots, such as "detection time" and "response time".

The "Civil Aviation Authority (CAA)" report also stated that the correct reaction time will be between 0.8 and 3.9 seconds depending on the flight phase. When the helicopter pilots give their full attention to the test tasks, the reaction times are very fast. However, they cannot always focus their attention on a single job.

Generally, the average detection time varies between 1 and 2.5 seconds (CAA, 1999). The helicopters have an average flight speed of 135 knots (250 km/h) in addition to the necessity of low flight in special missions such as military operations. Based on this data, a helicopter with a speed of 135 knots per hour travels about 200 meters in an average of 3 seconds. Therefore, an electrical wire detection system to be developed based on computer vision or image processing should be able to detect electrical wires at a distance of at least 200 meters.

1.2. Related Work

In this section, we will discuss the work done to recognize the electrical cables. Several methods have been proposed in the literature for the detection of electrical lines.

1.2.1. Academic literature

H. Saito and his colleagues proposed a laser-based method. The advantages of this laser system have been reported to be compact and operational. However, the lasers used in this system are not suitable for human eye health and have some weaknesses such as being affected by weather conditions (humidity, fog, heat, etc.) very quickly (Saito, 1992).

K. Sarabandi and his colleagues have recommended using Millimeter-Wave Radars to avoid a reduction in laser performance due to bad weather conditions. The periodic effects of the helical surface properties of the electrical lines at high frequencies (34 and 94 GHz) are important factors in the scattering behavior of electric lines. Therefore, due to their surface properties, radar back scattering occurs in the diagonal and horizontal polarized component, and these components use a unique separator signature (Sarabandi, 1999, Wei, 2008).

Again, K. Sarabandi and his colleagues in another study identified electrical cables using polarimetric SAR images obtained from the scattering of electromagnetic waves using the helical structure of powerlines (Sarabandi, 1994). They also proposed a method of extracting electric lines by using SAR images from Millimeter-Wave Radars operating at 35 GHz (Sarabandi, 2000).

Hiroyuki Yamaguchi and his colleagues proposed the use of the Millimeter-Wave Radar based on Bragg scattering method to determine the azimuth angular profile matching (APM) structure. As a result of the analysis at 35 GHz and 94 GHz frequencies, the probability of detection at 94 GHz was found to be higher (Yamaguchi, 2000).

Qirong Ma and his colleagues first applied the Hough transformation to detect all the lines in the millimeter wave radar image. To eliminate lines that do not belong to the electric lines, they used a 14-dimensional feature vector using Bragg patterns for each line they found and classed them using SVM classifiers. Finally, they determined the electrical wires by looking at the correlation between consecutive frames (Ma, 2011).

Another method is the detection of electrical lines based on the linear characteristics of the powerline, and therefore these methods are called line-based methods. Guangjian Yan and his colleagues first identified the line with a dash mask they identified. Secondly, they used the Radon transformation to express each line as a vector according to the slope and length criteria, and finally, using the Kalman Filter, they filled the gaps between the lines to obtain more robust lines (Yan, 2007).

Joshua Candamo and his colleagues first analyzed consecutive frames in the video and found the changing pixels. They then used the Canny edge detector to map edge areas of the varying pixel areas (property map) and used morphological filters (8 linked components) to eliminate unbound pixels (Candamo, 2006; 2008; 2009).

Again, in a similar study by Joshua Candamo and colleagues, they first performed edge detection to identify candidate pixels. They then used an 8-way chain-code histogram method to eliminate noise detected as an edge and to detect straight lines. Then, in order to analyze whether the detected lines were powerlines, they determined the Gaussian distributions of the pixels in a given region around each line and eliminated those that were not in the same group (Candamo, 2010).

Zhengrong Li and colleagues have developed a PCNN (Pulse-Coupled Neural Network) to separate the powerlines from the background. They then applied the Hough Transform to find the powerlines in the background separated image (Li, 2008, 2010).

Yuee Liu and colleagues used a directional filter to detect the differences in the image in the first stage and then used region magnification, bound component analysis and line addition methods to obtain lines from the resulting tread points, respectively (Liu, 2012).

Ömer Emre Yetgin and his colleagues have shown that electrical line detection can be done with high accuracy and fast line finding method such as EDLines, LSD and Hough Transform (Yetgin, 2013; 2015).

Xiaoyan Luo and his friends have proposed a method using new images created by combining RGB and NIR images. Then, electrical wire detection is performed by applying edge detection, line detection and verification methods in new fusion images (Luo, 2014).

Biqin Song et al. found primarily FDOG values of gray level images. They then applied a morphological filter to the FDOG response to obtain an edge map and using edge maps, they obtained line pools. Then, they realized the validation process with the developed graphical cutting model (Song, 2014).

Jun Zhang and his colleagues proposed a system of two stages (training and re-cognition). Poles and wires were determined during the training phase and the spatial correlation was calculated between them. A Bayesian probability classification model based on the correlation values was trained. During the detection phase, the detected posts and wires were classified using a pre-trained model (Zhang, 2014).

Carol Martinez and her colleagues have developed a system based on the prototypes of the power poles in their video frames. First, vertical components were found by applying a vertical Sobel filter and Hough Transform on each frame. Then, a specific area in the vertical components was divided into windows and the feature vector was obtained by using a HOG identifier in each window. Finally, the resulting feature vectors were classified using MLP (Multilayer Perceptron) (Martinez, 2014, Shan, 2015).

Feng Tian and his colleagues applied bilateral filtering to improve their video and then applied the Hough Transform to detect electrical wires (Tian, 2015).

Chaofeng Pan and his colleagues suggested a 3-step method. In the first step, they have extracted an edge map from each image using the directional filters. The extracted edge map has both electrical lines and background edges. In the second step, they used a CNN structure consisting of 5 hidden layers and 2 classes to eliminate the effects of the background (there is a powerline, there is no powerline). They determined the configuration parameters of CNN according to the fixed trial and error in the training phase and trained the CNN network. When the pictures in the database they used were examined, it was found that the pictures were taken from the ground. In other words, they used images with a background sky. In the last step, line detection was performed using Hough Transform (Pan, 2016). In reality, images with background shapes should be used in the background. For this reason, in our works, real images taken from the plane, images with ground shapes in the background were used (Yetgin, 2017).

For powerline detection, various systems have been developed to assist pilots. These systems will be briefly discussed below.

1.2.2. Industrial application literature

1.2.2.1. Powerline detector

An electronic unit for the limited detection of power cables has been developed (Vengalattore, 2008). This detector detects the electromagnetic field around the electrical wires and gives a warning to the pilot that it is close to an electric wire and consists of an electronic unit in the cockpit and a whip antenna mounted on the helicopter surface. This system can detect an electromagnetic field at a distance of about 1800 meters depending on the distance between the powerlines and the current (power) it contains. The warning sound increases as you approach the powerline. The electronic unit operates with 28 Vdc from the aircraft and can detect electrical wires at a frequency of 60 Hz. This detector can produce healthy results only for active powerlines and fails for other cables, inactive lines, phone lines, etc. (Vengalattore, 2008).

1.2.2.2. WSPS (Wire Strike Protection System)

Powerline accidents produce serious results for helicopters. For example, if the helicopter hits a cable with enough speed, the electrical wire may break the helicopter glass. Protection systems protect the helicopter from the wires that hit the windshield. Another commercially available system, namely WSPS, consists of one or more cutters mounted on the helicopter surface. The aim is not the detection of the cable, but rather elimination of the cable wire before it may impose a threat. There is a deflector vertically in the middle of the windshield and this deflector allows the electric cable to slide from the helicopter surface to the cutter.

In order for the cutters to be effective, the flight speed of the helicopter must be higher than 30 knots. According to the information obtained from the manufacturer, if a helicopter with a WSPS is struck by an angle less than 60°, there is a possibility that WSPS may not be able to cut its electrical wire. However, if a wire has hit the cutters on the helicopter, this cutter must be replaced with a new one (Vengalattore, 2008).

1.2.2.3. OASYS (Obstacle Avoidance Laser Radar System)

OASYS is a 35 GHz radio-frequency radar mounted on the nose of the helicopter and is used to detect obstacles in the direction of flight. This radar constantly scans obstacles in the field of view. At any time, the system uses the data from the aircraft GPS to identify the flight path of the aircraft and the three potentially risky areas in this direction. If an object is in any of these 3 zones, a warning is given to the pilot according to the proximity of the obstacle and the distance of the object to the aircraft and its location on a screen in the cockpit is indicated. The system scans a large area around the helicopter during take-off, but the scanning area narrows as the helicopter moves at high speed. The OASYS has a maximum coverage of 1600 meters, an optimal coverage of 800 meters and a minimum coverage of 2 meters. The system can also work effectively in rainy and foggy weather. The OASYS cannot detect objects above or behind the aircraft (Vengalattore, 2008).

1.2.2.4. LOAS (Laser Obstacle Awareness System)

LOAS is a warning system that can detect small objects such as wires. It uses a laser that is not harmful to the eye. The system can be directed to the desired direction by the pilot. The minimum coverage of the system is 2 m and the maximum coverage is 2 km. The viewing angle of the scan head can vary between 180° horizontally and between "+ 30° and -90°" in the vertical (Vengalattore, 2008).

1.2.2.5. HELLAS (Helicopter Laser Radar System)

HELLAS is a system that can detect as thin objects as wire. The system uses a laser that is not harmful to the eye and can be mounted anywhere on the helicopter. The air space around the helicopter is shown on the screen and the obstacles in the flight path of the helicopter are marked. HELLAS is a system that helps pilots in low visibility conditions. The coverage area of the system is reduced in bad weather (Vengalattore, 2008).

2. EXPERIMENTAL WORK

2.1. Aerial Image Dataset

Creating a database of aerial images taken from an aircraft is a difficult and very expensive method. Therefore, the limited number of aerial images has been used for the powerline detections in the literature. The aerial image dataset used in this study was created by TEIAS. The dataset contains VL and IR videos for the same scene. The airborne image system is shown in Figure 2.1.

(a) External appearance of the helicopter.   *(b) Imaging system mounted on a gimbal.*

Figure 2.1. *Airborne imaging system implemented by TEIAS (Yetgin, 2017).*

The video resolutions are 576 × 325 for IR and 1920 × 1080 for VL. As can be seen from the Figure below, an image database consisting of the pictures we obtained from the videos was created.

The first data set consists of 6000 VL and 6000 IR images in 128 × 128 dimensions (Figure 2.2.). Each image spectrum contains 4000 negative and 2000 positive samples. Also, the pictures in the database are classified as "Easy" and "Difficult", according to their contents. The videos were obtained from different geographical locations in Turkey (21 different cities). The samples were chosen to provide various difficulties due to different backgrounds, lighting and weather conditions.
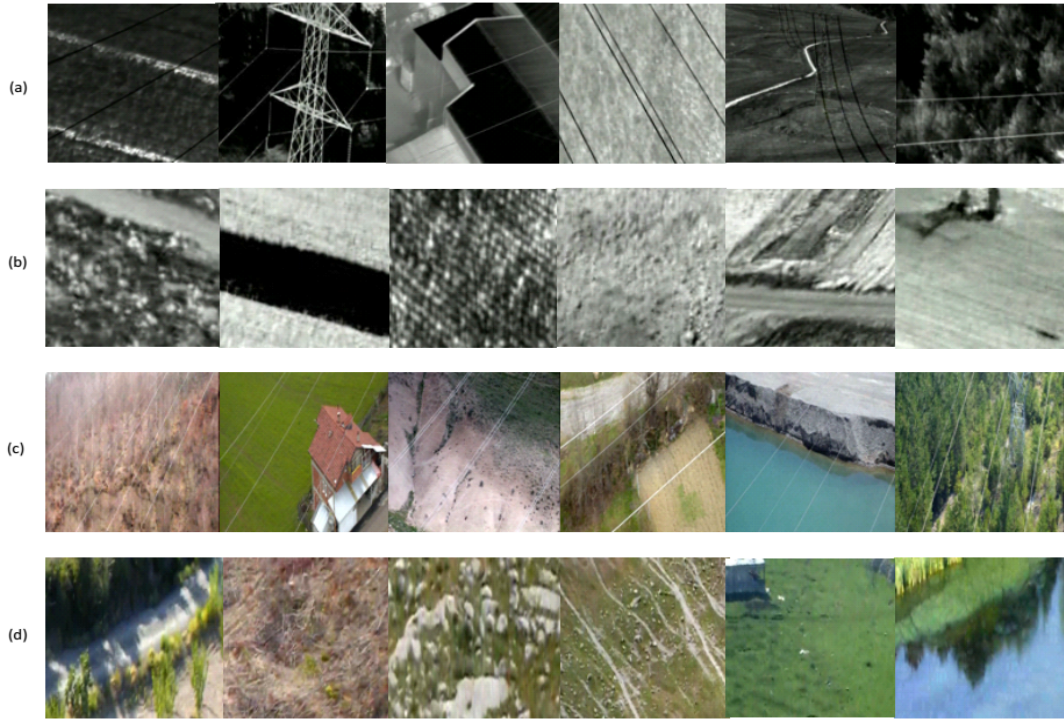
Figure 2.2. *Examples from the aerial image dataset. (a) IR images (with powerlines), (b) IR image (without powerlines), (c) VL images (with powerlines) and (d) VL images (without powerlines).*

The second data set consists of 400 VL and 400 IR images with 512×512 dimensions. Each image spectrum consists of 200 negative and 200 positive sample images and their Ground Truths. Electrical lines are generally warmer or colder than the environment, which significantly increases the visibility of electrical wires in IR images. However, this effect is reversed when the line temperature is close to the ambient temperature. When creating a database, we aimed to select examples of IR images that represent both of these conditions. In addition to proposing a method for powerline recognition, an important contribution to this study is to provide a real database for the problem of powerline recognition that can be used to compare studies in the literature. The database can be downloaded from (Yetgin, 2017).

Due to the limited amount of useful load handling of small aircraft such as heli-copters, computer vision systems offer several advantages for applications such as low size, weight and power consumption, wire detection. In order to assist pilots in aircraft, there are imaging systems using various equipment, such as thermal (night) and night vision cameras. Each of these imaging systems has its own specific purpose. When pilots are difficult to fly, they try to fly safely using these imaging systems.

Within the scope of this thesis, studies have been carried out on adequate remote recognition of electrical wires by using methods not previously used in the literature. In addition, a database of images taken from the aircraft not yet found in the literature was created.

DCT / FFT, LBP, and HOG methods are commonly used to extract features in image / pattern recognition applications. CNN is a machine learning method and it can be used as both feature extractor and classifier. In our study, the aforementioned methods have been used for detection powerlines in scene together with the various classifier methods.

2.2. Image Feature Extraction

2.2.1. Discrete cosine transform (DCT)

A signal / picture is subdivide to the sub-bands in the frequency domain with using DCT. The information in a signal / picture is scattered across various frequencies in the frequency domain. By using DCT, it is revealed how much density the information in the picture is in the frequency component. A signal/picture provides a conversion from a spatial domain to a frequency domain with using DCT, and in this respect is a similar conversion to a Discrete Fourier Transform (DFT).

Both Discrete Fourier Transform (DFT) and Discrete Cosine Transform (DCT) perform the following similar functions: Both represent a discrete time vector of finite length as a sum of the basic functions scaled and shifted. The difference between the two is the basic type of function used by each conversion. While DCT uses only (real-valued) cosine functions, the DFT uses a series of complex exponential functions with a series of harmonics.

DCT is often used in lossy data compression applications such as JPEG image format. The DCT representation of a signal, compared to other transformations such as DFT, is that most of the energy of the signal / image causes it to condense in a small number of coefficients. This is one of the most desirable features for a compression algorithm. If the original (time or spatial area) signal can be approximately represented using a relatively small set of DCT coefficients, we can reduce the data storage requirement by storing only significant amounts of energy-consuming DCT outputs.

2.2.1.1. Advantages and disadvantages of DCT

DCT is orthogonal (the opposite is equal to transpose and maintains its energy) and is a linear transformation and can be calculated with rapid algorithms. Energy conservation means that the sum of the squares of the values of the pixels in the image is equal to the sum of the squares of the coefficients obtained after DCT. Therefore, it is understood that there is no loss of information as a result of DCT.

The advantage of DCT is that it can collect the energy in the original picture to several frequency coefficients, depending on the correlation between the pixel values in the image. So many DCT coefficients have zero or very small values. Orthogonality means that the signal's auto-correlation is low. If the auto-correlation of a signal is high, that is, if there is a strong relationship between the signal values or if there is a strong correlation between the pixels in the image (the correlation is high), many DCT coefficients will be zero or too small.

The first DCT coefficients refer to low frequency DCT coefficients. High-value DCT coefficients are called high-frequency components, and for the 2-D matrices (images), the number of components with higher frequencies to the bottom right increases. For pictures, regions with low frequency components represent slow transitions / variations in the picture, and regions with high frequency components represent the location of rapid transitions / changes in the picture. Therefore, low-frequency components contain a lot of information about the overall image, while high-frequency components include the edge of the image.

If there is a pattern in the image, it can be said that the correlation between the components in that image is high. In the case of high correlations, the values of the low frequency components are high and the values of the high frequency components are low. For images without any pattern, no information can be obtained about which frequency component will receive low values. They may all be high or low value.

Inverse DCT (Inverse DCT) provides the original data from DCT coefficients. This conversion can occur 100% lossless depending on the data. In addition, if IDCT is applied to the signal which is obtained by making other coefficients zero excepts the first few coefficients on the DCT coefficients, the difference of obtained signal with original signal (mean square error) is very low. This is because the majority of the energy of the original signal is collected in the first few DCT coefficients. This means that the signal has high correlation.

2.2.1.2. Dynamic range adjustment (LOG-DCT)

When the magnitudes of coefficients have a huge variation, it may be difficult to depict small differences due to the exaggerated dynamic range. A popular method to suppress the dynamic range whilst retaining visualization of the differences at small values is taking the logarithm of the values. "LOG-DCT" is the logarithm of the DCT coefficients of an image. The logarithm function is often used for two purposes in many image processing / pattern recognition applications. The first is to minimize the changes in the illumination that is usually encountered in the preprocessing phase. For this purpose, the logarithm of the DCT is taken and the problems caused by the illumination are minimized.

Second, due to any conversion the pixel values between 0-255 in the image can be very large and a wide range. In this case, trading with these values increases both the workload and the interpretation of the data. For this reason, logarithms of these large values are obtained and the data is formed to smaller values. Thus, Although the process load resulting from the size of the data is reduced, the larger range covered by the data is narrowed to provide better interpretable data.

For example, the DCT of the IR image, which has pixel values between "0-214", takes values between -5095.3-26036 in both a high value and a wide range, while the logarithm of the absolute value of these values (log(abs(dct))), "-10,4236-10,1672" (Figure 2.3).

| (a) | (b) | (c) |
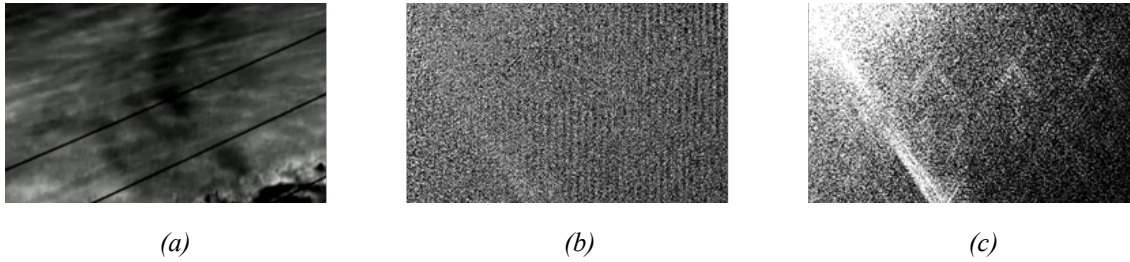
Figure 2.3. *LOG-DCT examples for IR image. (a) IR image (with powerlines), (b) DCT of (a), (c) LOG-DCT of (a).*

For example, the DCT of the VL image, which has a pixel value of "56-240", takes values between "-1089.2-59730" in both a high value and a very wide range, while the logarithm of the absolute value (log (abs (dct))), "-13,1476-10,9976" (Figure 2.4).

| (a) | (b) | (c) |
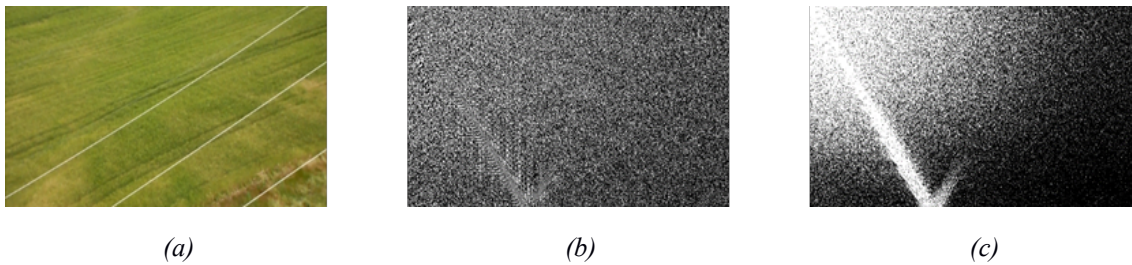
Figure 2.4. *LOG-DCT examples for VL image. (a) VL image (with powerlines), (b) DCT of (a), (c) LOG-DCT of (a).*

2.2.1.3. Formulation

For an exemplary matrix MxN in image A, Type-II 2-D DCT (commonly used) is defined as follows:

$$B_{pq} = \alpha_p \alpha_q \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} A_{mn} \cos\left[\frac{\pi(2m+1)p}{2M}\right] \cos\left[\frac{\pi(2n+1)q}{2N}\right] for \begin{array}{l} 0 \leq p \leq M-1 \\ 0 \leq q \leq N-1 \end{array}$$

$$\alpha_p = \left\{ \begin{array}{l} \sqrt{1/M} \quad p = 0 \\ \sqrt{2/M} \quad 1 \leq p \leq M-1 \end{array} \right\} \qquad \alpha_q = \left\{ \begin{array}{l} \sqrt{1/N} \quad q = 0 \\ \sqrt{2/N} \quad 1 \leq q \leq M-1 \end{array} \right\} \qquad (2.1)$$

Where p and q represent vertical and horizontal frequencies, respectively. It has not yet been determined whether the DCT of electrical wires is a global structure or a detail. As can be seen from the Figure 2.5, a different pattern is observed in the DCT of the lines in various angles.
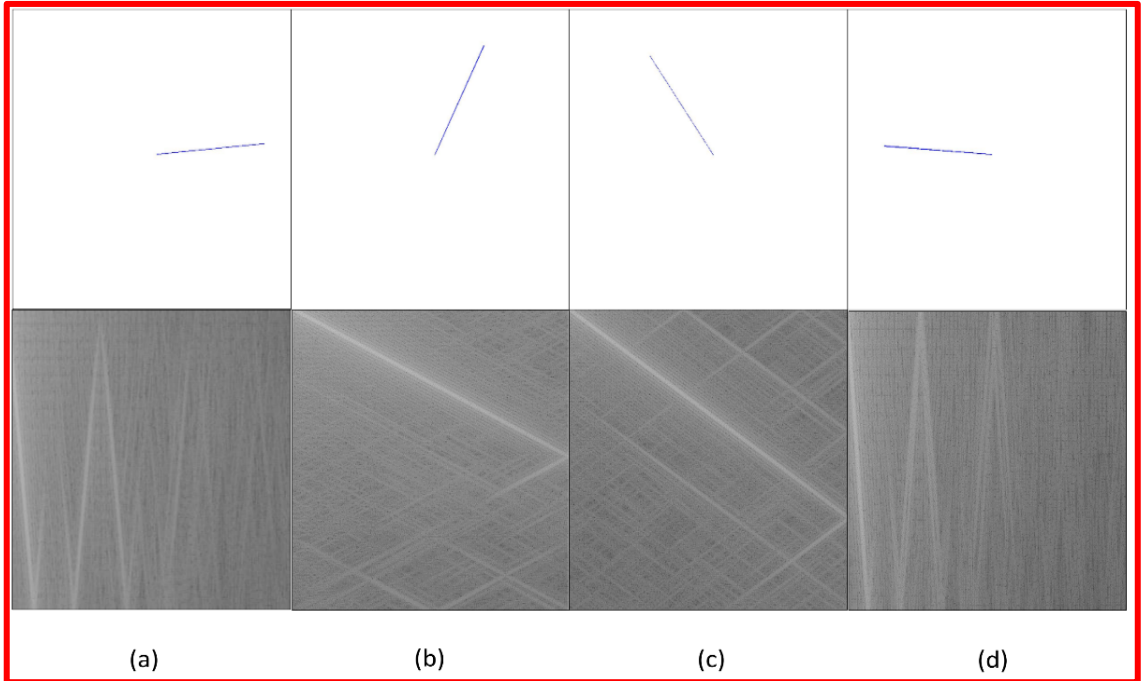
Figure 2.5. *Line image and its abs-log DCT (a) A 5º line image and its abs-log DCT, (b) A 63º line image and its abs-log DCT, (c) A 126º line image and its abs-log DCT, (d) A 176º line image and its abs-log DCT.*

16

Unfortunately, as in the picture Figure 2.6., electrical wires do not always have a simple background. Therefore, the real problem is to look for electrical wires in the real images. Otherwise, Hough Transform (HT) (Duda, 1972) would be an excellent and clear choice to detect artificial lines on images, but it was observed that actual images had a very negative impact on the performance of HT. Therefore, more robust results will be obtained with DCT. The Figure 2.6. shows IR and VL images with and without electrical wires and corresponding abs-log DCT images.
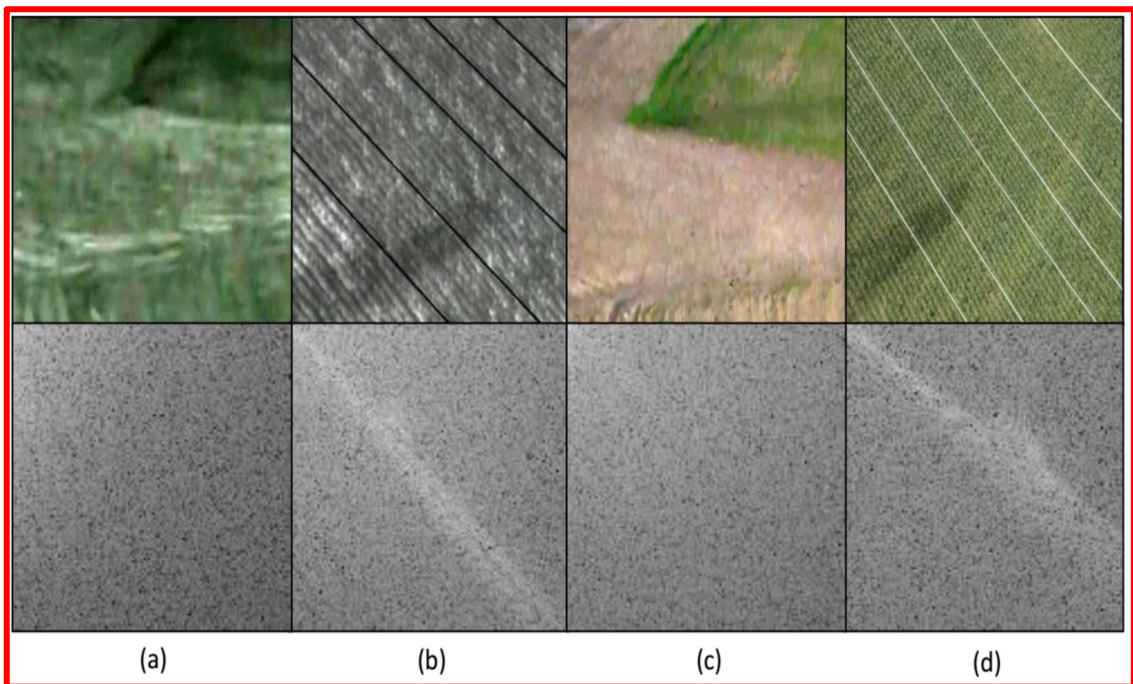
Figure 2.6. *IR and VL images and its abs-log DCT (a) IR image without powerlines and corresponding abs-log DCT, (b) IR image with powerlines and corresponding abs-log DCT, (c) VL image without powerlines and corresponding abs-log DCT, (d) VL image with powerlines and corresponding abs-log DCT.*

2.2.1.4. DCT feature selection strategies

Because most of the signal energy is present at low frequencies, the choice of features in the upper left of the DCT is a conventional approach. In this study, we considered the selection of the property area in the 2x2, 4x4, 8x8, 16x16, 32x32, 64x64 and 128x128 DCT window sizes (as shown in Figure 2.7 (a)) except the DC coefficient itself. In our experimental studies, we defined this path of DCT feature selection as "Classical Selection (CS)".

In contrast to the classical approach, electricity lines are expected to be at higher frequencies. Therefore, a trial worth approach is to select the DCT field starting from the highest frequency region (as in Figure 2.7 (b)) of similar variable field dimensions. In our experimental studies, we defined this path of DCT feature selection as "Reversed Selection (RS)".

Another alternative to the above two strategies is to select block regions on the DCT data in the form of a pavement, as seen in Figures 2.7 (c, d and e). In our experimental studies, we defined this path of DCT feature selection as "Proposed-2 (Patch-Based-PB)". For all DCT block selection methods above, the number of properties increases as the block size increases. This is expected to improve classification performance. However, excessive adaptation to the model may cause degradation of performance during the testing phase. This effect was also investigated in the following experimental sections.
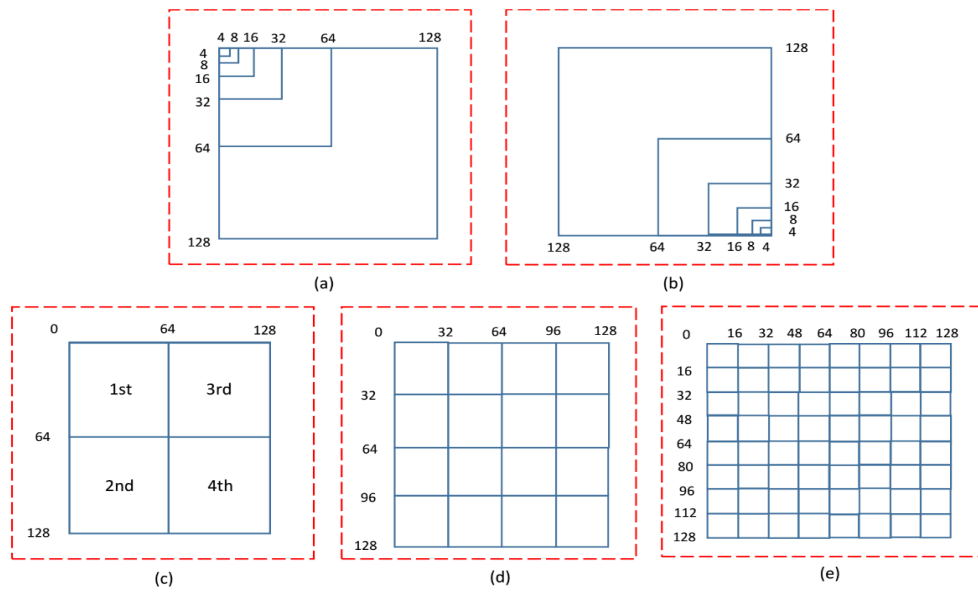
Figure 2.7. *DCT features selection methods (a) Low frequency favored selection of DCT features at increasing sizes, (b) High frequency favored selection of DCT features at increasing sizes (Proposed-1), (c-d-e) Patch-Based Selection of DCT features at different sizes (Proposed-2)*

2.2.2. LBP and HOG

LBP is widely used as a feature extraction method in the areas of image processing and computer vision (Ojala, 2002). It provides resistance to different lighting conditions and is simple in terms of calculation. Quite simply, the 8-neighbor with a pixel of interest is compared. As a result of the comparison, instead of the neighboring pixel, when the neighboring pixel is larger than the compared pixel, a value consisting of 1 and otherwise a value of 0 is written, resulting in the 8-bit code.

The decimal value of this 8-bit number naturally identifies local structural information around the corresponding pixel. Another very popular object detection method is HOG (Dalal, 2005). It is known that the gradient information obtained from the image carries distinctive information between the foreground and background in the image. In the HOG method, the image is divided into blocks and the gradients in each cell are calculated (at certain angular intervals - usually around 20 degrees). HOG deals with the histogram of the gradient angles, and these gradient angles create a distinctive property for the objects.

These two methods and the properties obtained from DCT were tested / used on the same classifiers to detect images containing electrical wires.

2.2.3. Experimental Results for Features

The properties obtained using the above-described feature extraction / selection methods were fed into the Naïve Bayes, Random Forest and SVM classifiers to define scenes with electrical wires. Classification success was evaluated according to accuracy rate (%).

Figure 2.8 shows the accuracy of Naive Bayes, Random Forest and SVM classifier results in different dimensions obtained by the LBP method for (a) IR and (b) VL images.

The corresponding results for the HOG method are shown in Figure 2.9. The results for DCT properties begin with Figure 2.10, where the property dimensions are selected according to CS. Figure 2.11 illustrates the results for DCT feature selection this time with an inverse selection path (Proposed-1 as described in Figure 2.7).

A quick comparison between Figures 2.11 and 2.12 shows that selecting DCT samples from high frequency regions (Figure 2.12) significantly improves the electrical wire detection performance (Proposed-2 as described in Sec.2.2.1.4).

By comparison, LBP performs better on VL images, while HOG on IR images performs better. Nevertheless, with the patch-based selection, DCT-based classification gives good results from both.
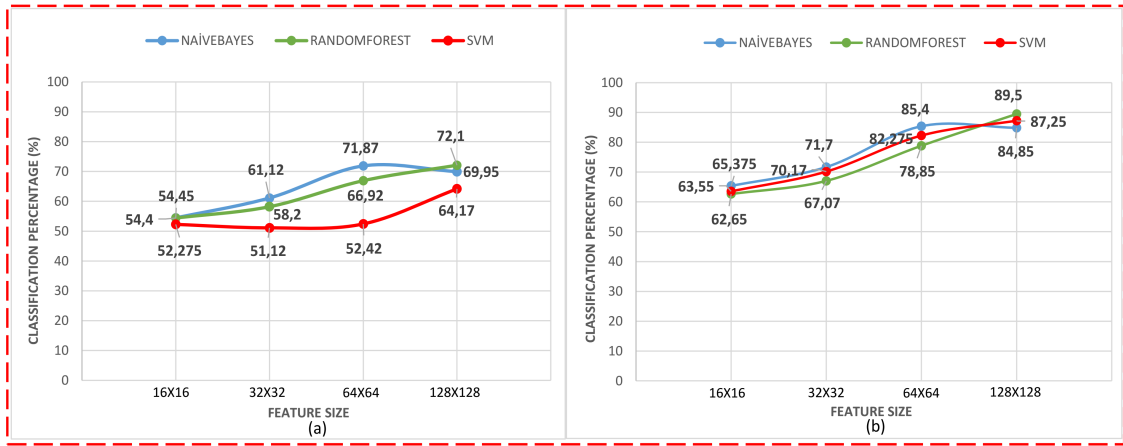
Figure 2.8. *Classification results using LBP at various feature sizes for (a) IR and (b) VL images.*
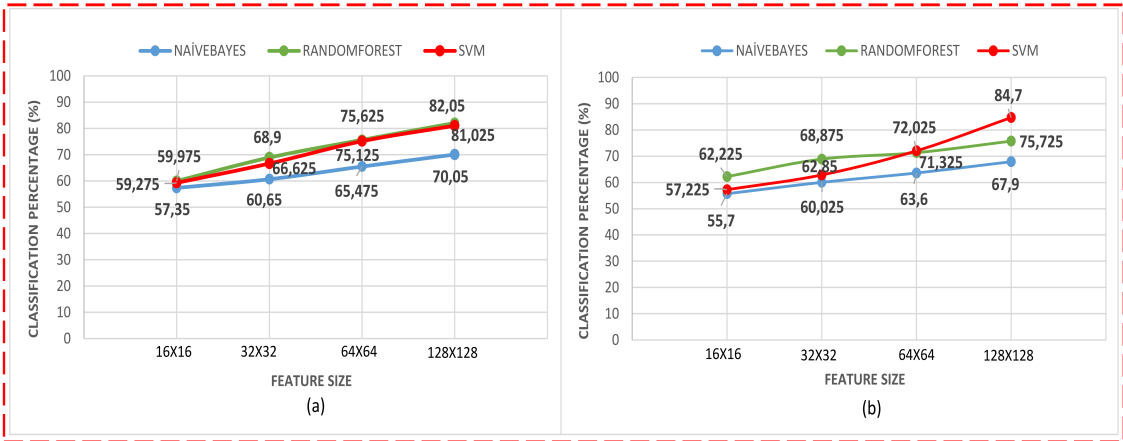
Figure 2.9. *Classification results using HOG at various feature sizes for (a) IR and (b) VL images.*
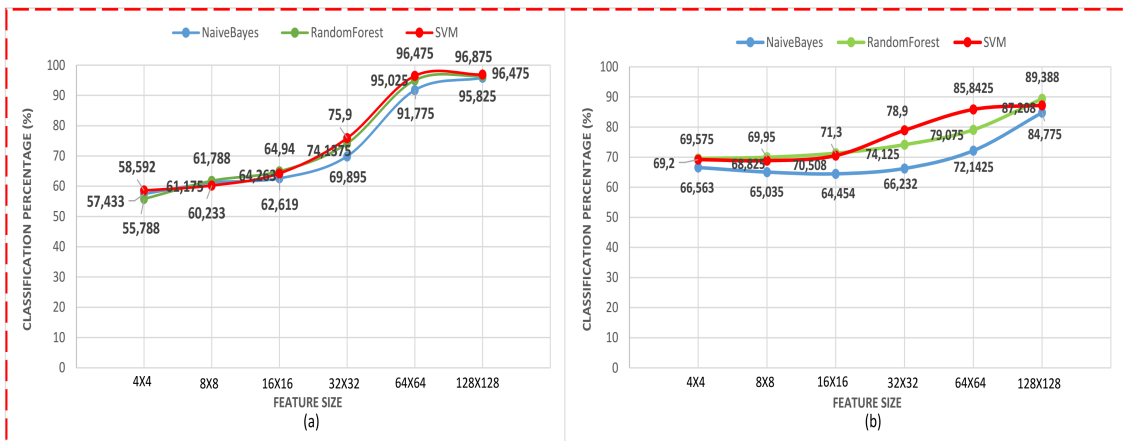
Figure 2.10. *Classification results using CS at various DCT feature sizes for (a) IR and (b) VL images.*
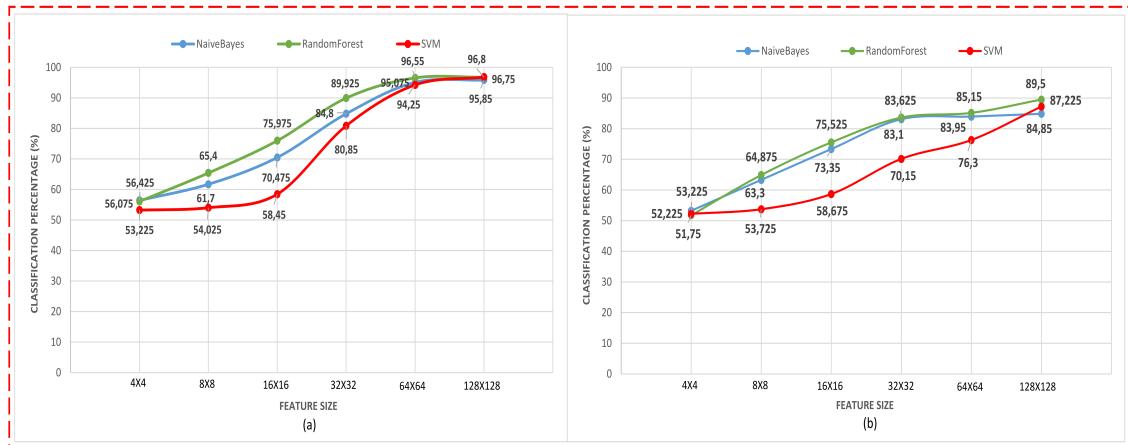
Figure 2.11. *Classification results using Proposed-1 (RS) at various DCT feature sizes for (a) IR and (b) VL images.*

The final comparison is performed with the-Patch-Based Based DCT Selection, which is called Proposed-2. In this method, DCT samples from low and high frequency components are selected. As we can see in Figure 2.12 (a), while the Patch-Based Based DCT Selection is the most successful method for IR images, Reverse DCT Selection method is the most successful method for VL images.
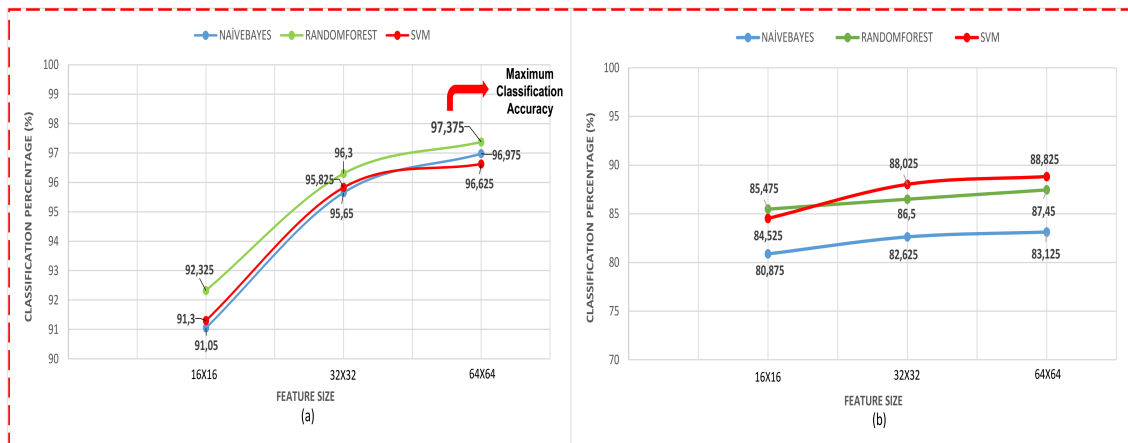
Figure 2.12. *Classification results using Proposed-2 (PB) Selection at various DCT feature sizes for (a) IR and (b) VL images.*

In the Classical and Proposed-2 methods, the conversion of matrices into 1-D property vectors before classification was carried out by zigzag scanning of DCT blocks. On the other hand, in the Proposed-1 method, a 1-D feature vector was created by zig-zag scanning method (from the highest frequencies to low frequencies) starting from the bottom right of the DCT matrix.

Table 2.1 and 2.3. show the maximum results from the IR and VL image database, IR and VL results are also presented as confusion matrices (Tables 2.2 and 2.4) to better explain false alarm and missing rates. Typically, LBP and HOG are not available in 4x4 and 8x8 property sizes.

Table 2.1. *Detailed results (at maximum accuracy, corresponding to Patch-Based-64x64 (Proposed-1) selection and Random Forest classifier) for IR database.*

Class	TP Rate	FP Rate	Precision	Recall	F-Measure
1 (with cable)	0,991	0,044	0,958	0,991	0,974
0 (without cable)	0,957	0,009	0,991	0,957	0,973
Weighted Avg.	0,974	0,026	0,974	0,974	0,974

Table 2.2. *The confusion matrix of the IR database.*

A	b	Classified as
1982	18	a = 1 (**with cable**)
87	1913	b = 0 (**without cable**)

Table 2.3. *Detailed results (at maximum accuracy, corresponding to Reverse-128x128 (Proposed-2) selection and Random Forest Classifier) for VL database.*

Class	TP Rate	FP Rate	Precision	Recall	F-Measure
1 (with cable)	0,930	0,140	0,870	0,930	0,899
0 (without cable)	0,861	0,071	0,924	0,861	0,891
Weighted Avg.	0,895	0,105	0,897	0,895	0,895

(a)   (b)
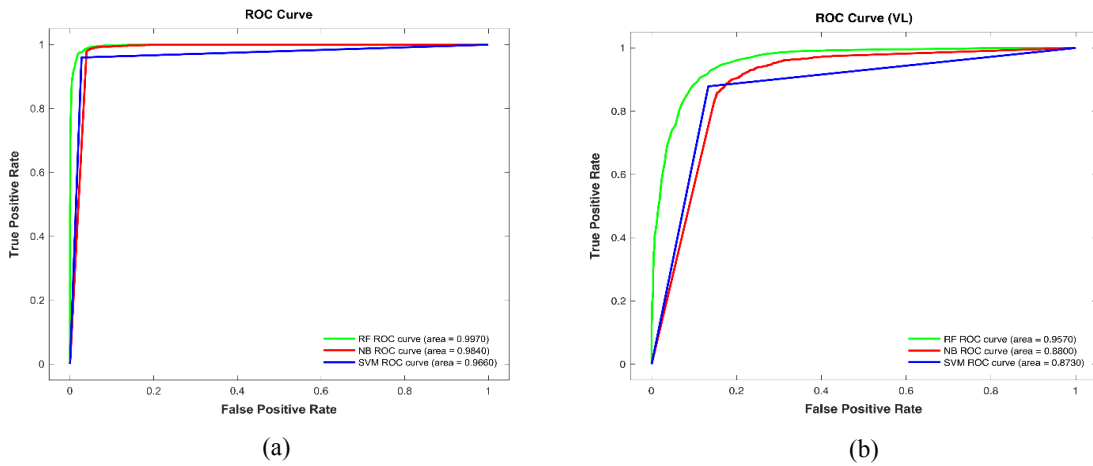
Figure 2.13. *Receiver Operating Characteristic (ROC) curves for (a) Patch-Based Feature Selection Method for IR and (b) Reverse Feature Selection Method for VL images with different classifiers.*

Table 2.4. *The confusion matrix of the VL database.*

a	b	Classified as
1859	141	a = 1 (**with cable**)
279	1721	b = 0 (**without cable**)

For the corresponding maximum accuracy conditions presented in Table 2.1-2.3, receiver operating characteristic curves are also determined and are presented in Figure 2.13. When these curves are examined, it can be seen that most of the classifiers provide a very efficient (sharply curved) ROC curve with the most efficient classifier, Random Forest (labeled green and labeled as RF). Again, the IR images show higher performance (Figure 2.12 (a)). In particular, the effects of objects such as roads, buildings, railways in the background of the images on the recognition performance of electrical wires appear to be less in IR images than in VL images. The success and robustness of the classification results of the proposed algorithms are illustrated in more detail by comparing with popular methods based on edge / line detection throughout Figures 2.14-2.19.

The Figures presented in Figures 2.14 and 2.19 illustrate the case in which the images with electrical wires are properly classified by our method and the edge / line detection methods successfully find the electrical wires (despite the scattered backgrounds).
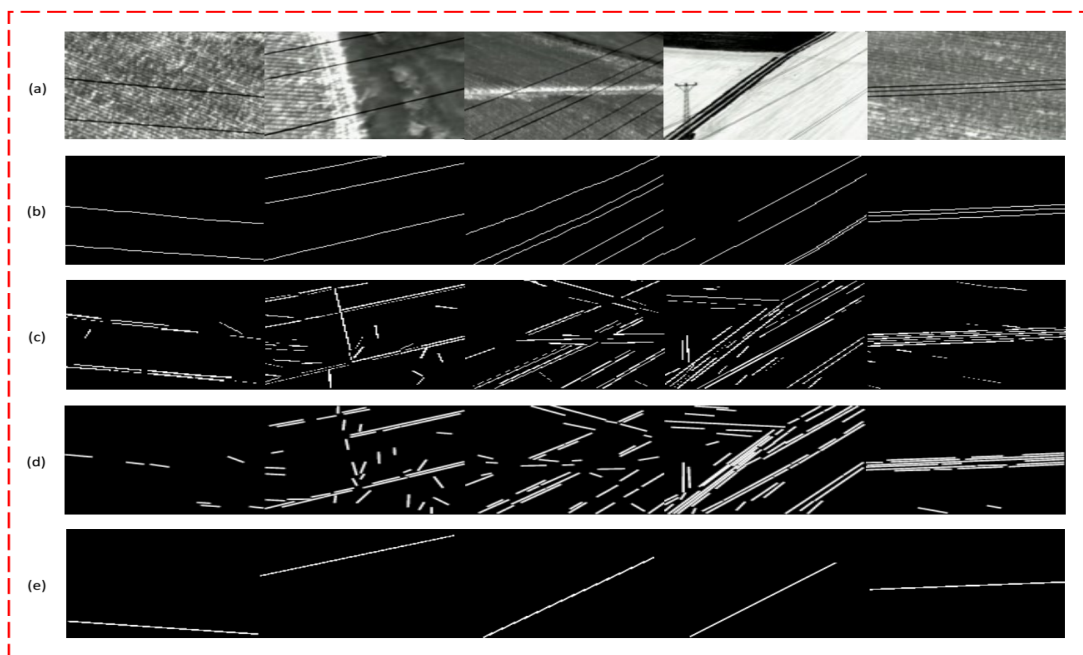
Figure 2.14. *Edge/Line detection results : (a) IR images (with powerlines, correctly detected by our method), (b) Ground Truth of (a), (c) EDLines, (d) LSD and (e) HT. All methods are successful.*
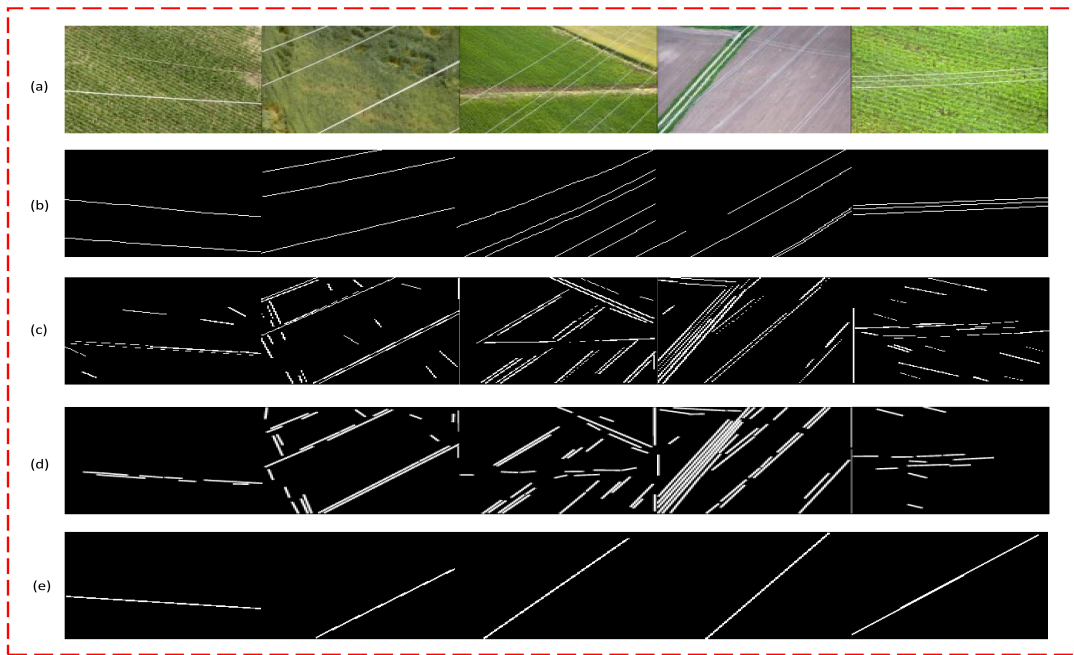
Figure 2.15. *Edge/Line detection results : (a) VL images (with powerlines, correctly detected by our method), (b) Ground Truth of (a), (c) EDLines, (d) LSD and (e) HT. All methods are successful.*

Figures 2.14 and 2.15 illustrate situations in which the image includes electrical wires and is properly classified by our methods, but edge / line detection methods usually detect edges / lines corresponding to the wrong lines/edges.
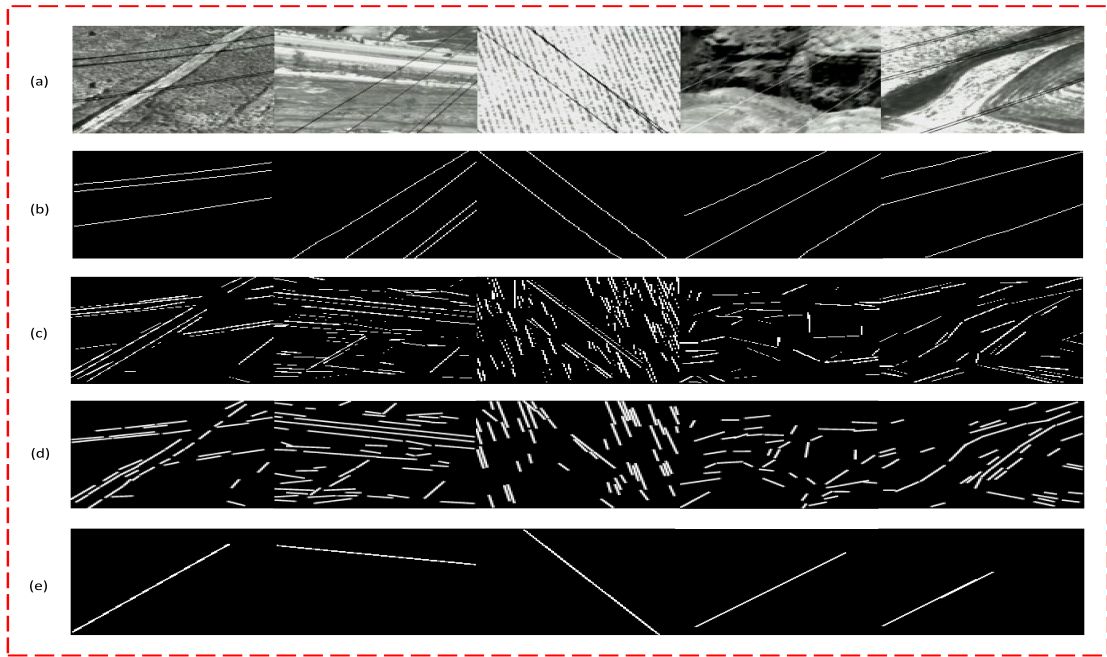
Figure 2.16. *Edge/Line detection results : (a) IR images (with powerlines, correctly detected by our method), (b) Ground Truth of (a), (c) EDLines (d) LSD and (e) HT. Edge/line detection methods incorrect.*
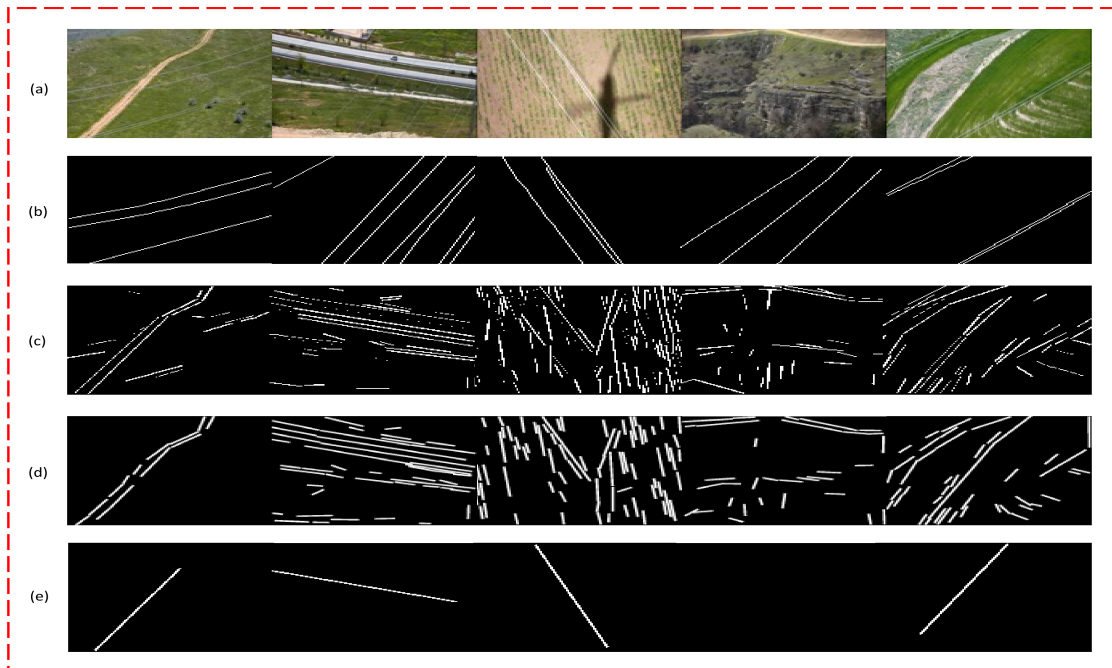
Figure 2.17. *Edge/Line detection results : (a) VL images (with powerlines, correctly detected by our method), (b) Ground Truth of (a), (c) EDLines (d) LSD and (e) HT. Edge/line detection methods incorrect.*

Although the images presented in Figures 2.16 and 2.17 are are correctly classified by our methods, but it has been observed that edge / line detection methods give more false alarms.
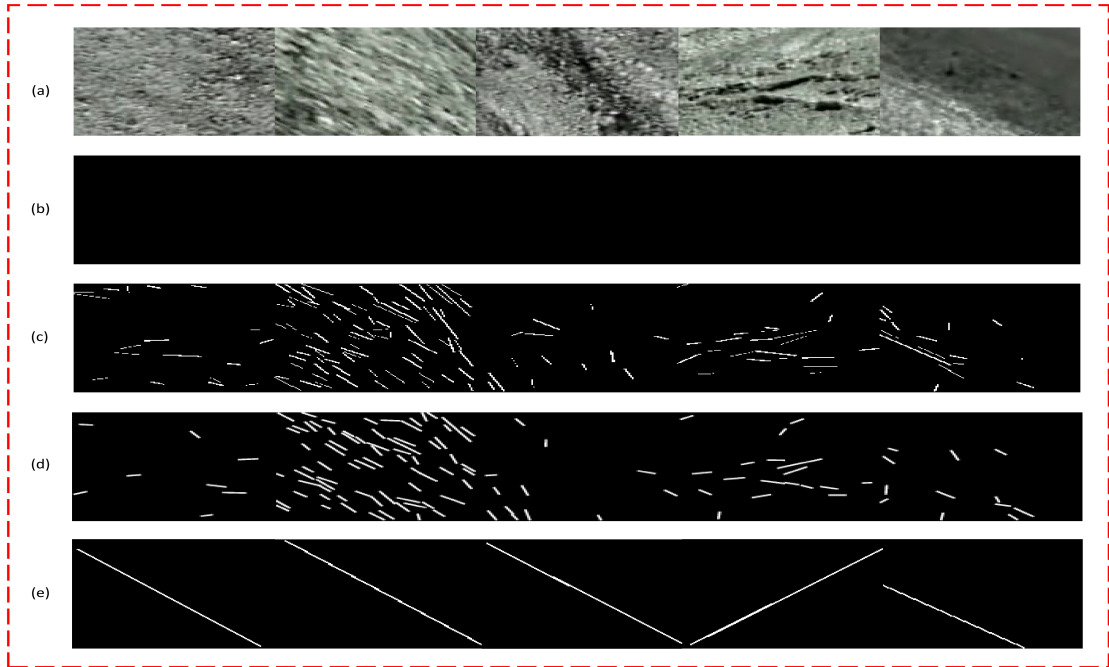
Figure 2.18. *Edge/Line detection results: (a) IR images (without powerlines, correctly identified by our method), (b) Ground Truth of (a), (c) EDLines (d) LSD and (e) HT. Edge/line detection methods incorrect.*

Figure 2.19. *Edge/Line detection results: (a) VL images (without powerlines, correctly identified by our method), (b) Ground Truth of (a), (c) EDLines, (d) LSD and (e) HT. Edge/line detection methods incorrect.*

The detailed results for the various DCT feature selection methods (with LBP and HOG), applied to various classification methods, are presented in Table 2.1 (for IR) and Table 2.3 (for VL images).

The best performing examples of each class are highlighted in bold. The experimental results are presented for a small number of data samples (ie 4x4 or 8x8), for this case, the conventional DCT selection method provides more successful results (compared to higher frequency features only).

However, starting from the 16x16 dimension, the larger and higher frequency features start to offer better performance, so the Proposed-2 and Recommended-1 methods provide more successful results than other methods. If a comparison is made between the classifiers, it can be said that Random Forest and SVM classifiers perform better than Naïve Bayes. Interestingly, IR images provide higher classification results in larger feature sizes. However, for small property sizes, this does not apply.

Considering the criticality of accurate detection of images with electrical wires in air vehicles, the Random Forest classifier applied to IR images and the choice of Patch-Based DCT (Recommended-2) with the size of 64x64 achieved the highest success with a test classification accuracy of 97.38%. Experimentally, the 64x64 size region corresponds to the 4th region we identified in the DCT matrix (corresponds to the DCT samples in the highest frequency range). This region is in fact the same as the 64x64 region in Proposed-1.

The only difference is the low-high frequency zig-zag sequence in Proposed-2 and the high-low-frequency (traditional) zig-zag scan in Proposed-1. Experimentally, the so-called proposed-2 method appears to perform better, particularly using Random Forest classification. Edge / line detection-based methods (as shown in Figure 2.14-2.19) exhibit very high false positive rates and their classification performance is usually below 55% for IR and below 48% for VL databases.

The Table 2.5 shows the IR and VL image classification performances. VL images were more successful for 4x4 and 8x8 features, but for larger feature sizes, IR images were more successful.

Table 2.5. *Comparison of IR and VL Database Classification Results.*

FEATURE EXTRACTION / SELECTION METHOD	FEATURE SIZE	SUCCESSFUL DATABASE	SUCCESSFUL CLASSIFIER
CLASSICAL SELECTION	4x4	VISIBLE	RANDOM FOREST (69,58%)
CLASSICAL SELECTION	8x8	VISIBLE	RANDOM FOREST (69,95%)
PATCH-BASED SELECTION (PROPOSED-2)	16x16	INFRARED	SVM (96,20%)
PATCH-BASED SELECTION (PROPOSED-2)	32x32	INFRARED	SVM (96,30%)
PATCH-BASED SELECTION (PROPOSED-2)	**64x64**	**INFRARED**	**RANDOM FOREST (97,38%)**
CLASSICAL SELECTION	128x128	INFRARED	SVM (96,88%)
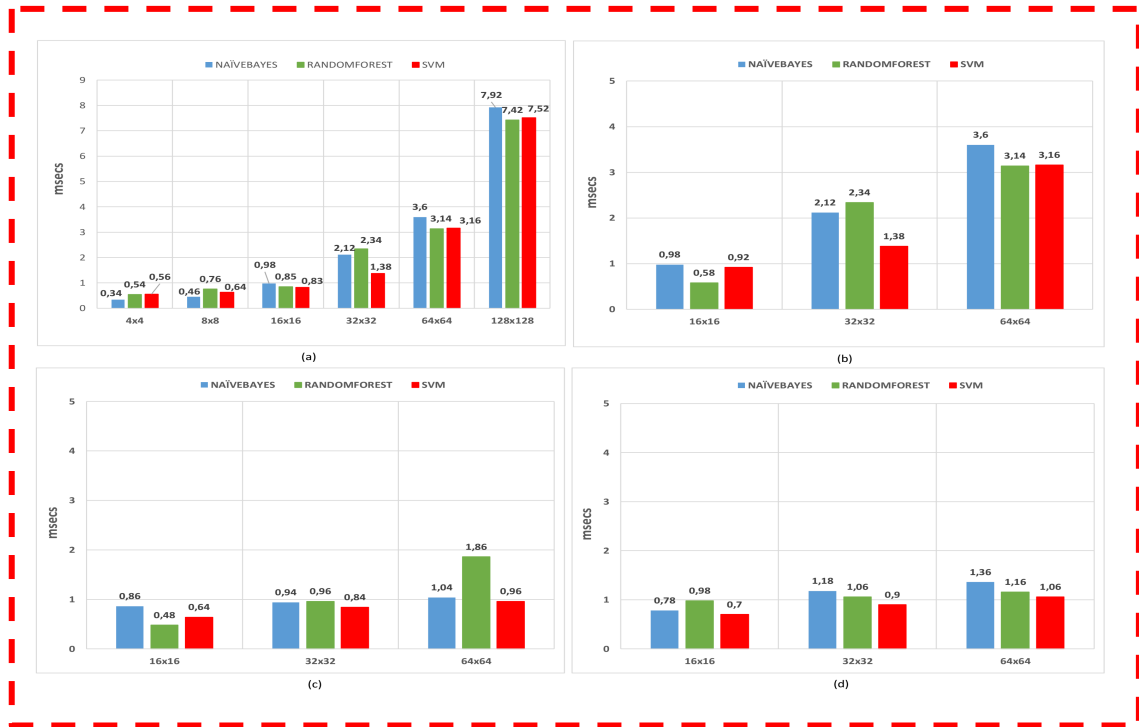
Figure 2.20. *Process time per image using: (a) Classic and Reversed Selection (Proposed-1) of DCT Feature and Classification Methods, (b) Patch-Based Selection (Proposed-2) of DCT Feature and Classification Method, (c) LBP Feature Extraction and Classification Method and (d) HOG Feature Extraction and Classification Method.*

As expected, the property size affects the computation time for all classifiers. The classification results of the LBP and HOG based classification results and the 16x16 DCT features work at approximately the same time. However, as the size of the DCT feature grows, the time consumption begins to increase. However, there is no significant difference between the working hours of the classifiers. The critical observation is that the detected operating times are below the limit of 3 seconds. Therefore, it can be said that the proposed methods can be easily implemented in real-time systems.

2.3. Classifiers

An important step in recognizing the image containing the electrical wire is to use the features on the classifiers. In our problem, there are 2 classes with and without electrical wires. Our aim is to determine whether there is powerline within the given image by classification methods.

As classifier, famous classifiers such as Naïve Bayes (Murphy, 2001), Random Forest (Breiman, 2001) and Support Vector Machines (SVM) (Cortes, 1995) were used. 10-fold method was used in the test phase of the classification process.

2.3.1 Naïve Bayes

The regular Bayesian classifier is based on obtaining a threshold for classifying two probability density functions with known conditional distributions. The Bayes classifier, which can also be described as a stationary Gaussian random distribution, uses two mean vectors and two covariance matrices to separate two different classes using quadratic hyper-planes.

For this problem, vector / class membership probabilities are calculated (counted) according to the given ground truth and a priori (hence a-posteriori) probability model is created. In the test phase, an observation feature vector is simply put into the probability model and a decision is made by maximizing the possibility of a-posteriori. Due to its simple and relatively successful performance, this classifier has been used for power-line classification problem (Murphy, 2001).

2.3.2. Random Forest

Random Forests begin from the entire dataset in the property field. Rather than creating a single decision tree, a number of decision trees are created by combining the best representative features of the tree, and these decision trees are combined. In this classifier, random feature extraction / selection plays an important role (Breiman, 2001).

2.3.3. Support Vector Machines (SVM)

SVM is a very popular classifier that aims to achieve an optimal hyperplane as a decision function on a higher dimensional space that is extended via support vectors and appropriate kernels. The closest feature vector of the hyperplanes for each class is called support vectors. These support vectors are then mapped to a vector space (possibly a higher dimensional) (with linear or non-linear kernels), so that the separation between the support vectors and the hypersensitive surfaces is as large as possible.

After completion of the training (using a linear SVM based on logistic regression), relevant decisions were obtained (Cortes, 1995).

2.4. Deep Learning and Convolutional Neural Network

The classical method for general purpose object recognition is to select the properties in an image and classify them by means of a classifier (Csurka, 2004). CNN architecture is a neural network that combines both property extraction and classification (He, 2016).

In recent years, large image data sets such as ImageNet (Krizhevsky, 2012) have resulted in the discovery that CNN models are very efficiently scaled with data. ImageNet, Large-Scale Visual Recognition Data Set (ILSVRC) (Donahue, 2014) is used by deep CNN models every year (Zeiler, 2014, Bengio, 2013, Razavian, 2014) and is known to be one of the best data sets for high-level recognition.

The use of the ImageNet database varies according to the purpose to be performed. If the data content of the target task is significantly different, no preliminary training is performed in ImageNet. For example, Liu et al. Instead of using a pre-trained ImageNet network to find the face region in the picture, they used a separate network, trained only with facial images, to define the facial features (Gao, 2017). In some cases, ImageNet pre-training is completely omitted if the target task field is substantially different (Penatti, 2015), (Hu, 2015).
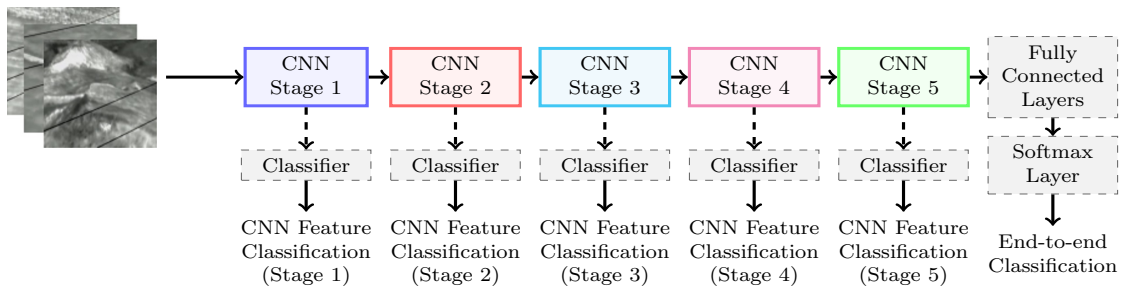
Figure 2.21. *Two alternative methods for using CNNs for powerline recognition : end-to-end classification and CNN feature classification.*

The image classification with a pre-trained and fine tuned ILSVRC is also called a transfer learning method. Deep networks benefit significantly from transfer learning (Yetgin, 2017). However, it is not right to expect that ImageNet pre-training is always useful. Qualifications learned by deep networks become more specialized as layers progress (Ioffe, 2015).

We propose two alternative methods for the use of CNNs in the context of the electrical wire recognition problem (see Figure 4.1). The first method is end-to-end classification. In this method, we start with a CNN (ILSVRC) designed to be used for image classification. Normally, this CNN has 1000 outputs for each class in the ILSVRC dataset, and finally a softmax layer. In our problem, we replace this final layer with a 2-layer softmax layer to make a binary classification (no wire or not). Then, we only train this last layer until we get convergence. Subsequently, fine tuning is carried out for feature extraction and classification sections.

In end-to-end classification, the entire network can be trained jointly. On the other hand, in CNN feature classification, the feature extractor and the classifier are trained separately. The disconnect is represented by dashed arrows.

In the second method, we used CNN to extract features on the images in the database. We split CNN to 5 stage for feature extraction. We obtained properties from each stage. The properties obtained were classified using 3 different classifiers. In particular, less GPU memory and processing power were spent than we used to feature specific areas of CNN only. In the following subsections, we will introduce certain CNN architectures and classifiers that we use to implement these methods.

2.4.1. CNN Architectures
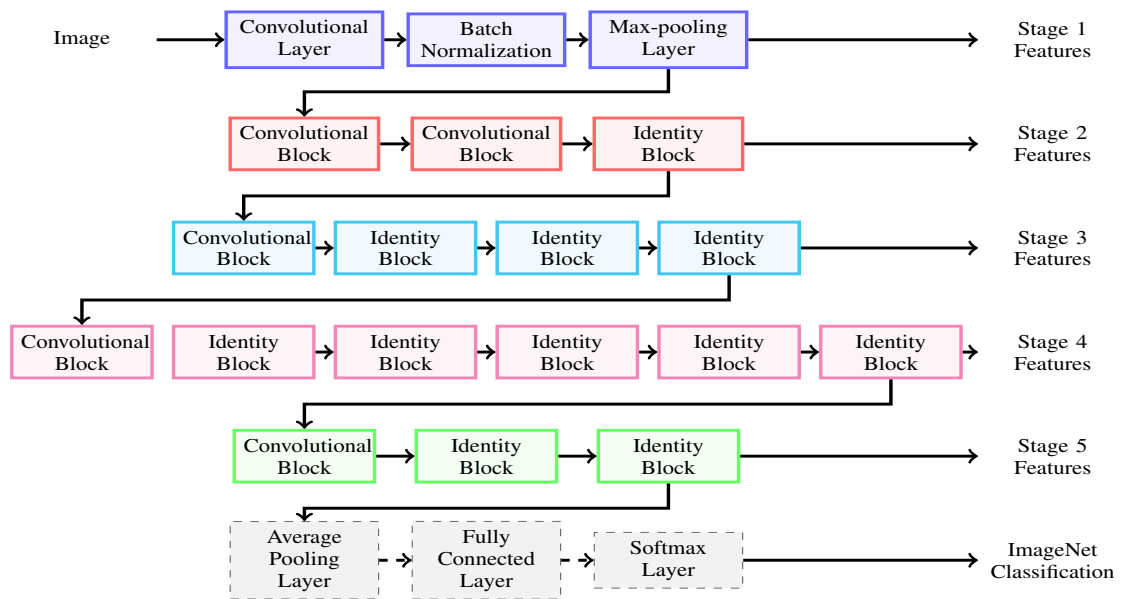
(a) VGG-19
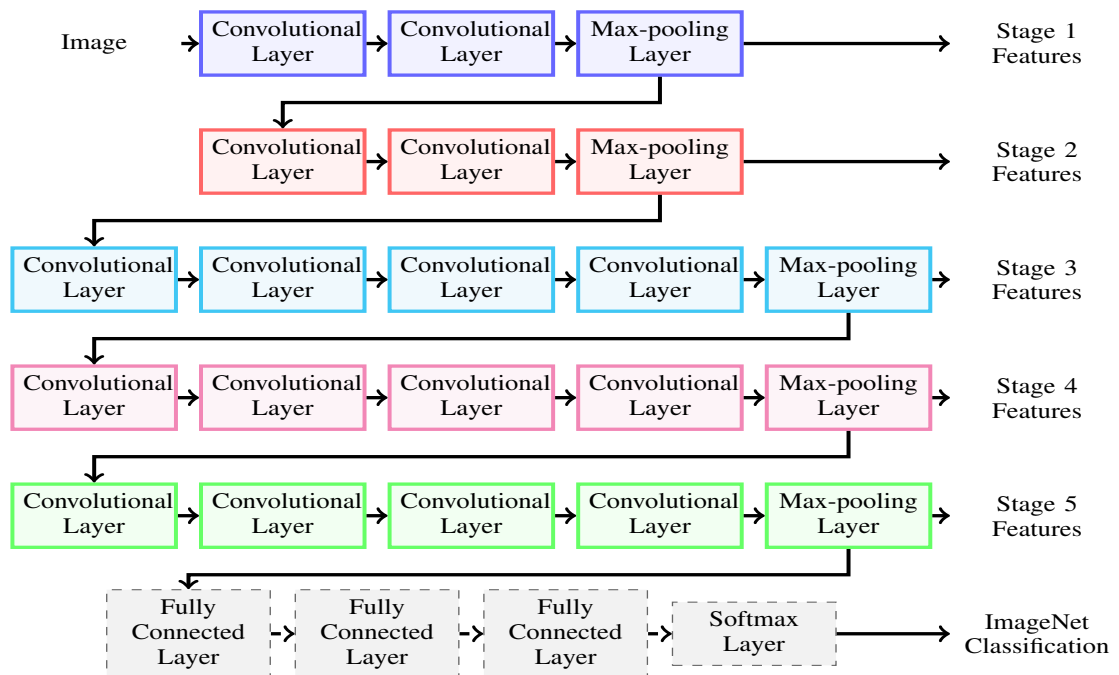
(b) ResNet-50

Figure 2.22. *CNN architectures used in the study. The convolutional stages of the architectures are illustrated in different colors. Convolutional and identity blocks of ResNet-50 are expanded on its left hand-side. (a) VGG-19 and (b) ResNet-50.*

In this section, we will briefly introduce the CNN architectures that we use with the recommended method, VGG-19 and ResNet-50. We chose these two architectures (ResNet-50 and VGG-19) because of the successes on the ILSVRC image database.

2.4.2. VGG-19

In previous architectures, convolution has manually adjusted the core dimensions for each layer. The main contribution of the VGG model, using the 3×3 fixed kernel size, performs as well as the architectures using larger kernel sizes, thus greatly simplifying the architectural design process.

2.4.3. ResNet-50

Krizhevsky and others first simplified the architectural design by designing Inception Blockes and using them to create GoogLeNet. Similarly, ResNet was designed as a combination of these blocks. The main difference of ResNet is that it is much deeper than previous architectures (Krizhevsky, 2012).

2.4.4. Experimental Works

In this section, we present the experimental results for the two recommended methods. End-to-end classification using CNN and classification by feature extraction.

2.4.5. Implementation Details

We ran all experiments with 10-fold cross validation. The data set for each fold is divided into 70% training data, 20% validation data, and 10% test data. The learning rate for the CNN final layer was started at 0.1 and reduced to five times when the loss of verification stopped. The fine-tuned learning rate was set to the same, but was started at 0.01. Weight reduction is set to 0.001 for all layers. The classifier parameters are kept as default values in Weka 3.8.

Two popular alternatives for image preprocessing have been tested. In the first, the average extraction method was used. In this method, the average pixel value in the image is subtracted from all pixel values to make images approximately zero-mean. The second one is scaling. In this method, the pixel values of all images are divided into 255 to scale between 0 and 1.

2.4.6. End-to-end Classification

Our experiments were initiated with pre-trained CNNs for ILSVRC image classification (hereinafter referred to as ImageNet pre-trained networks). These networks include filters, such as edge and block detectors, found in earlier layers that may be useful in most visual tasks.

As discussed earlier, the randomly initiated final layer is separately trained, followed by fine tuning of the entire network. The whole network has millions of parameters that can be fine-tuned. The number of free parameters increases the impressive power of the model. However, if there is no more training data, this also causes the model to be overfit. Since our training set is relatively small, we must limit the number of free parameters in the model. Therefore, we only fine-tune the final convolution stage (Step 5 in Figure 2.21) and limit it to the following layers.

See Table 2.6 for results where pre-trained networks are fed with mean-extracted images. By training only the last layer, we can see that we have achieved remarkable performance with the features learned before the training. ResNet-50 performs significantly better than VGG-19. Following this, we fine-tune the final stage of the networks. Here, we see that the ResNet-50 performance is further improved and results in the best performance reported in this study.

On the other hand, there was no improvement in the IR performance of the VGG-19. Confusion matrices in which step 5 is precisely adjusted are given in Table 2.7. Errors in the IR data set are balanced.

Table 2.6. *Classification errors in percentages for the end-to-end classification method (pre-training, mean subtraction preprocessing).*

	IR	VL	IR	VL
VGG-19	1.85	8.167	4.967	8.167
ResNet-50	0.65	1.0	**0.25**	**0.267**
	Trained last layer		**Trained last layer & Fine-tuned Stage 5**	

Table 2.7. *Confusion matrices for the end-to-end classification method (trained final layer and fine-tuned Stage 5, ImageNet pre-training, mean subtraction preprocessing). Rows are ground truths, and columns are predictions.*

ResNet-50						**VGG-19**					
IR	**Neg**	**Pos**	**VL**	**Neg**	**Pos**	**IR**	**Neg**	**Pos**	**VL**	**Neg**	**Pos**
False	3994	6	**False**	3994	6	**False**	3836	164	**False**	3691	309
True	9	1991	**True**	10	1990	**True**	134	1866	**True**	181	1819

False positive images had linear properties that misaligned the network. In the case of VL images, powerlines are not visible and false negative results have increased. Also, in most of these faulty pictures, we see that the electrical wires are close to the edges of the image. The reason for this is that CNNs have difficulty in recognizing environmental objects because of the first convolutional layers. This problem can be prevented by filling or transforming images.

In the second experiment, we used the scaling method instead of the average subtraction method. See the results in Table 2.8 and compare with Table 2.6. In this method, the average extraction is done for the pre-training process. The interesting thing is that after fine-tuning the 5th stage, the performance is significantly increased and comparable to the average extraction process. CNN architectures are naturally frequency selective. Therefore, Gabor is sensitive even to wavelet-like structures when they generate random weights. In other words, ImageNet pre-training may not have a significant effect on achieving the performance in Tables 2.8 and 2.6.

Table 2.8. *Classification errors in percentages for the end-to-end classification method (pre-training, scaling preprocessing).*

	IR	VL	IR	VL
VGG-19	57.183	44.433	1.467	1.967
ResNet-50	0.55	4.7	**0.217**	**0.55**
	Trained last layer		**Trained last layer & Fine-tuned Stage 5**	

To test the effect of ImageNet pre-training, we created networks with random weights using Xavier initialization and repeated the experiments. See Table 2.9 for results. It is clear that ImageNet pre-training is particularly useful for powerline recognition even when images are in the IR spectrum. However, untrained networks also performed better than random estimates.

Table 2.9. *Classification errors in percentages for the end-to-end classification method (random weights, scaling preprocessing).*

	IR	VL	IR	VL
VGG-19	22.283	21.633	27.933	**17.5**
ResNet-50	40.517	48.983	**13.333**	20.5
	Trained last layer		**Trained last layer & Fine-tuned Stage 5**	

The best performances in Tables 2.9-2.10 can be compared to the best performances in Table 2.9; this is equivalent to subtracting the properties of Stage 5 and classifying it with a single-layer network. Refer to Figure 2.26 for the receiver operator characteristic curve of the best performance tests in this section.

Table 2.10. *Detailed results (at maximum accuracy with using SVM classifier) for IR database with using ResNet-50/mean subtraction preprocessing (for Stage-4).*

Class	TP Rate	FP Rate	Precision	Recall	F-Measure
1 (with cable)	0,997	0,002	0,998	0,997	0,998
0 (without cable)	0,999	0,003	0,997	0,999	0,998
Weighted Avg.	0,998	0,002	0,998	0,998	0,998

Table 2.11. *Detailed results (at maximum accuracy with using SVM classifier) for VL database with using ResNet-50/mean subtraction preprocessing (for Stage-3).*

Class	TP Rate	FP Rate	Precision	Recall	F-Measure
1 (with cable)	0,992	0,014	0,986	0,992	0,989
0 (without cable)	0,986	0,008	0,992	0,986	0,989
Weighted Avg.	0,989	0,011	0,989	0,989	0,989
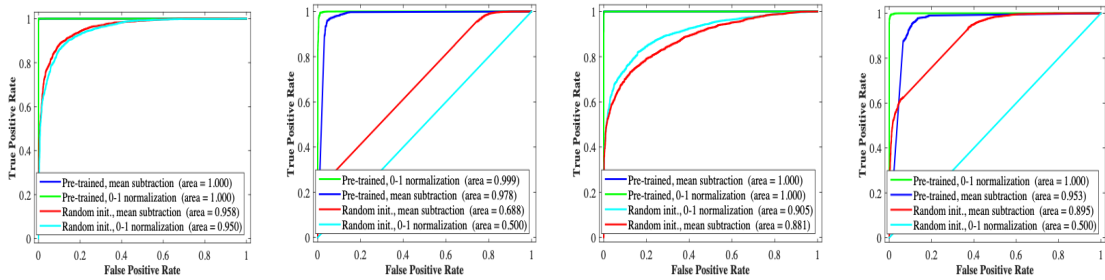
| (a) IR, ResNet-50 | (b) IR, VGG-19 | (c) VL, ResNet-50 | (d) VL, VGG-19 |

Figure 2.23. *Receiver operating characteristic (ROC) curves of the end-to-end classification method. The legend is in the order of decreasing area under the curve. Note that the curves of the best per forming methods overlap and occlude each other in some Figures.*

Table 2.12. *Classification errors in percentages for the CNN feature classification method (VGG-19, scaling preprocessing).*

	SVM	NB	RF	SVM	NB	RF
Stage 1	9.125	11.35	11.375	10.25	32.8	15.925
Stage 2	3.025	12.375	8.875	3.425	21.575	12.125
Stage 3	0.8	10.375	2.45	**2.275**	16.125	4.75
Stage 4	**0.6**	11.85	2.9	2.475	17.925	6.275
Stage 5	0.875	9.475	1.6	4.45	16.9	6.575
	IR			**VL**		

Table 2.13. *Classification errors in percentages for the CNN feature classification method (ResNet-50, scaling preprocessing).*

	SVM	NB	RF	SVM	NB	RF
Stage 1	21.9	7.875	7.25	36.725	20.975	**8.775**
Stage 2	11.2	12.925	6.275	29.025	29.45	10.125
Stage 3	12.45	21.15	6.05	30.175	35.575	14.925
Stage 4	11.6	23.025	**5.0**	27.875	35.2	13.425
Stage 5	10.2	17.325	5.1	25.95	35.525	14.675
	IR			**VL**		

Table 2.14. *Classification errors in percentages for the CNN feature classification method (VGG-19, mean subtraction preprocessing).*

	SVM	NB	RF	SVM	NB	RF
Stage 1	12.383	29.917	26.2	4.933	35.233	24.95
Stage 2	7.4	28.633	24.433	1.9	37.883	26.8
Stage 3	2.267	14.167	18.667	**1.15**	27.617	20.6
Stage 4	0.917	17.4	16.467	2.0	30.55	21.233
Stage 5	**0.85**	21.183	11.75	2.85	30.15	21.9
	IR			**VL**		

Table 2.15. *Classification errors in percentages for the CNN feature classification method (ResNet-50, mean subtraction preprocessing).*

	SVM	NB	RF	SVM	NB	RF
Stage 1	10.07	31.183	23.4	2.717	38.233	26.733
Stage 2	3.95	15.65	19.267	1.7	23.483	20.317
Stage 3	0.833	16.733	11.85	**0.883**	14.733	19.167
Stage 4	**0.417**	7.767	10.417	1.67	29.917	19.3
Stage 5	**0.417**	19.75	11.1	1.083	33.483	22.417
	IR			**VL**		

In addition, we see that random forests produce the best results from ResNet-50. See Figure 2.23 for the receiver operator characteristic curve of the best performance tests in this section. This suggests that the proposed method is particularly strong for mixing backgrounds (roads, buildings, railways, etc.) in the case of IR imaging. The proposed algorithm is further described by comparing the classification results with popular methods based on line detection along Figures 2.24 and 2.25.
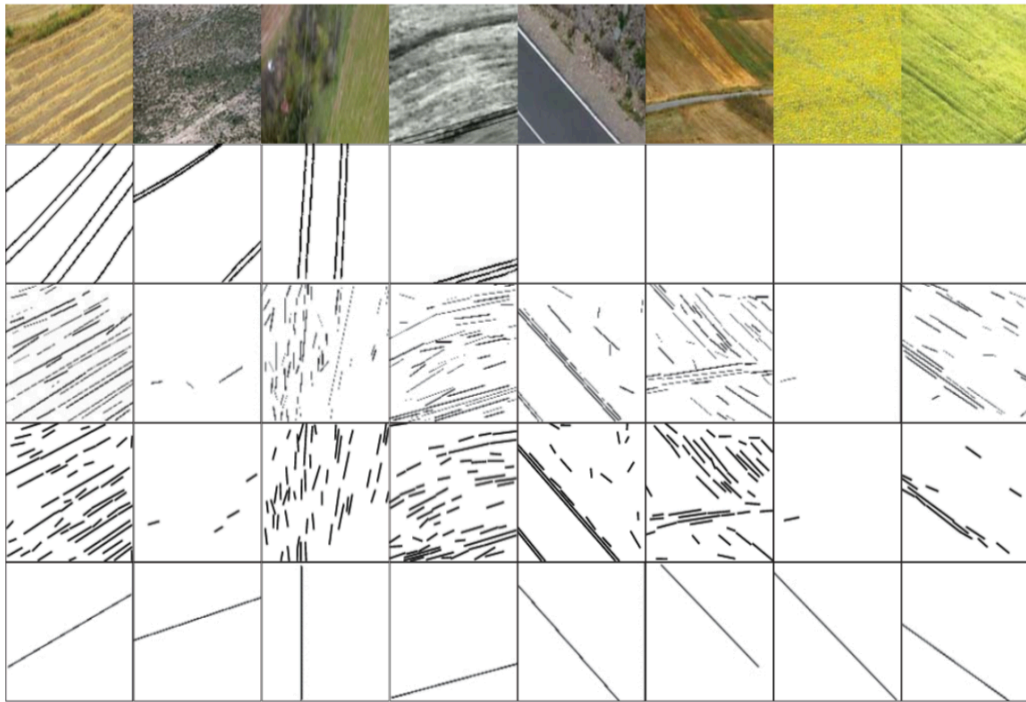

Figure 2.24. *The first row shows negative examples from both spectra, which were classified correctly by the end-to-end method.*
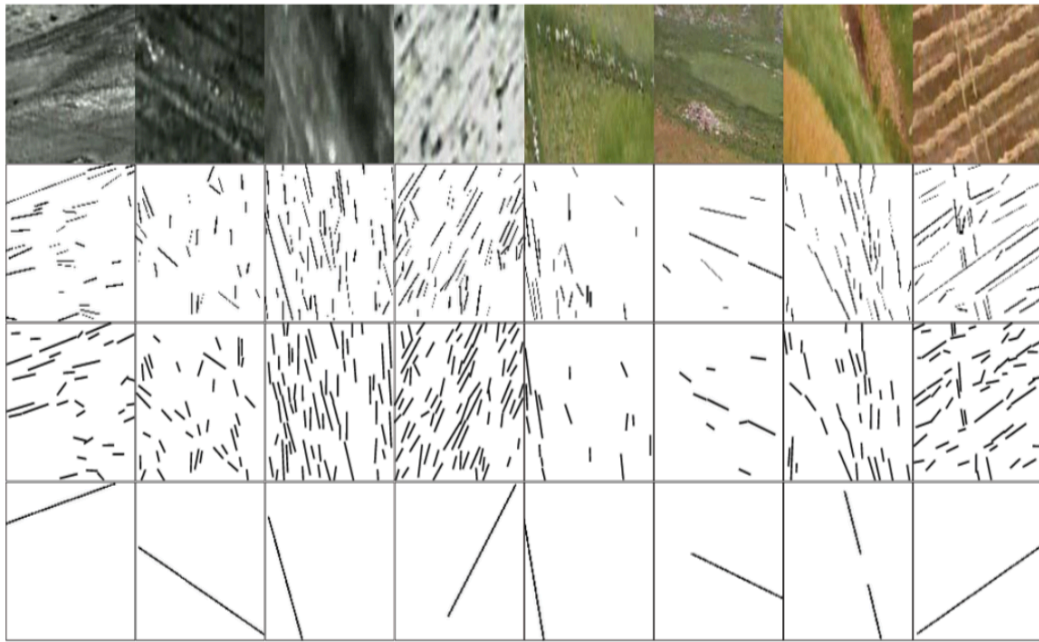
Figure 2.25. *The first row shows four negative and four positive examples, which were classified incorrectly by the end-to-end method.*

In general, line detection methods can be said to produce more false alarms. Existing line detection-based methods (as shown in Figure 2.24 and 2.25) exhibit very high false positive rates for IR and VL databases.

2.4.7. Running Time

The operating times given in Table 2.16 were obtained with an Nvidia GTX 1080 GPU and the operating times in Table 2.17 with the Intel Core i7 2.8 GHz CPU. Our best configuration, end-to-end classification with ResNet-50, runs at 21.7 ms, which is reasonable for a real-time application.

Table 2.16. *Running times of the classifiers for a single image in milliseconds.*

	Stage 1	Stage 2	Stage 3	Stage 4	Stage 5	End-to-end
VGG-19	3.7	4.1	5.4	8.0	10.0	10.6
ResNet-50	2.2	7.9	11.1	16.6	19.3	21.7

Table 2.17. *Cumulative running times of the CNN models for a single image in milliseconds.*

SVM	NB	RF
1.8	7.5	1.7

We see that we can improve working time by switching to VGG-19 and using intermediate stage features. For example, the VGG-19 feeds the Stage 4 properties to an SVM of 9.8 ms. However, this configuration offers approximately five times more errors.
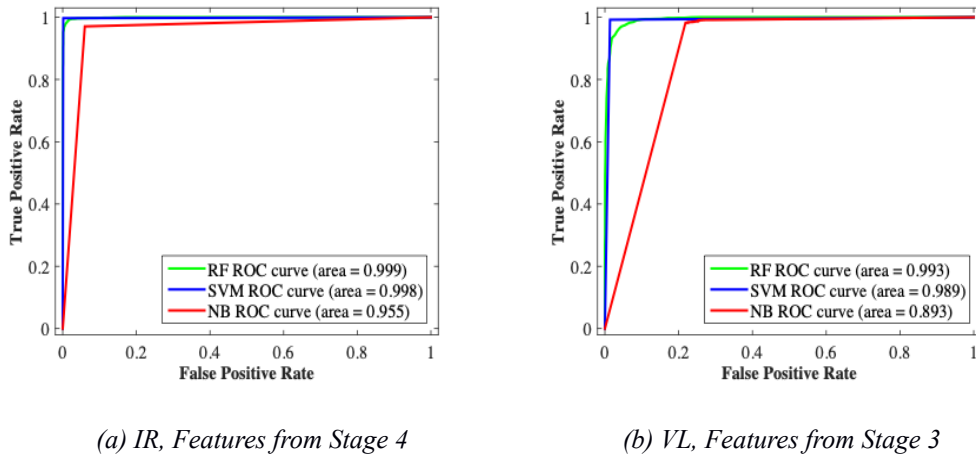

(a) IR, Features from Stage 4 *(b) VL, Features from Stage 3*

Figure 2.26. *Receiver operating characteristic (ROC) curves of the CNN feature classification method. ImageNet pre-trained ResNet-50 model is used with mean subtraction preprocessing. Stage 4 for IR and Stage 3 for VL are shown because they delivered the best results.*

2.4.8. Results

In this study, we proposed two CNN based powerline recognition methods for real time warning system. Both methods use CNNs designed for ImageNet object recognition. In the first method of end-to-end classification, the CNN was modified and co-trained for the target task.

In the second method, CNN is the property classification. In this method, the properties are subtracted from the intermediate stages of the CNN and fed to a classifier. Best results were obtained by end-to-end classification in which the network was previously trained with the ImageNet data set, and powerline images were pre-processed by the mean extraction.

The classification error was 0.025% for infrared images and 0.275% for visible light images. In almost all experiments, infrared images have been classified more successfully. This suggests that it is preferable to use infrared imaging for a powerline warning system. However, with visible light images the performance was also reasonable.

CNN property extraction has the advantage that it can use outputs from any stage of the network. The best results of CNN feature extraction were obtained by the properties obtained from the later stages which were close to the end-to-end classification results prior to fine-tuning.

The CNN feature classification allows the lower processing time to be balanced for a higher error rate. In the experiments without fine tuning, we observed that the preprocessing method was critical. Specifically, we should use the preprocessing method used in the preprocessing, which is the mean subtraction in our case. However, fine-tuning overrides the effect of the difference between pre-training and training methods used in education.

Although the architectures we used are designed for ImageNet object recognition, we have shown that they perform well in the target task. What's more, the architecture that performs better in the ImageNet object classification, performs better in powerline recognition, whether pre-trained or not. This implies that an architecture's ImageNet performance is a general indication of the performance of other visual tasks. Our data set consists of aerial images of visible light and infrared spectra. This is a very different domain than the ImageNet dataset. However, we observed that ImageNet pre-training was the most important positive factor in our experiments.

In CNN's stage analysis, maximum result is found in ResNet-50 (mean subtraction process) architecture, IR image dataset, false negative results in Figure 2.27.
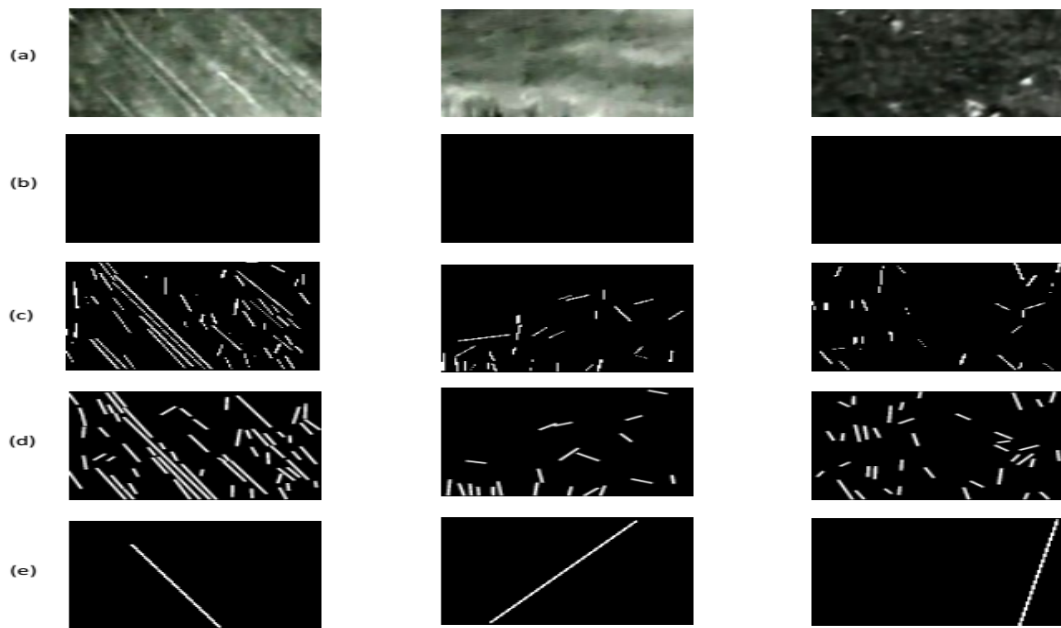
Figure 2.27. *In CNN's stage analysis, maximum result is found in ResNet-50 (mean subtraction process) architecture, IR image dataset, False positive results compared to other methods. Line detection results: (a) IR images (with powerlines, incorrectly detected with CNN stage analysis), (b) Ground Truth (c) EDLines, (d) LSD and (e) HT. Edge/line detection methods incorrect.*

3. CONCLUSION

In this thesis, it has been shown that powerlines can be detected using DCT and CNN methods at high accuracy and speed. The accuracies of both methods render them plausible alternatives for aircraft safety. When the two methods are further compared, it can be concluded that the CNN method gives higher accuracy results at the expense of a slower pre-processing stage. Conversely, the DCT method produces slightly less accurate results after a rapid preprocessing. However, it must be noted that both methods provide rapid and accurate detection results within the state-of-the-art aircraft safety regulations. Sequential and video based scene classification for detection of cable wires can be listed as possible future works.

REFERENCES

Bengio, Y., Aaron, C. and Pascal, V., (2013), "Representation learning: A review and new perspectives.", IEEE transactions on pattern analysis and machine intelligence 35.8, 1798- 1828.

Breiman, L., (2001), "Random forests.", Machine learning 45.1, 5-32.

Candamo, J., Kasturi, R., Goldgof, D., Sarkar, S. (2006), "Vision-based on-board collision avoidance system for aircraft navigation.", In Unmanned Systems Technology VIII, Vol. 6230, p. 62300X.

Candamo, J., Goldgof, D. (2008), "Wire detection in low-altitude, urban, and low-quality video frames.", In Pattern Recognition, ICPR, 19th International Conference on (pp. 1-4), IEEE.

Candamo, J., Kasturi, R., Goldgof, D., Sarkar, S. (2009), "Detection of thin lines using low-quality video from low-altitude aircraft in urban settings.", IEEE Transactions on aerospace and electronic systems, 45(3).

Candamo, J., Goldgof, D., Kasturi, R., Godavarthy, S. (2010), "Detecting wires in cluttered urban scenes using a Gaussian model", In Pattern Recognition (ICPR), 20th International Conference on (pp. 432-435), IEEE.

Chen, Z., Li, T., Zhang, X., Li, S., Hou, A., Liu, Y. (2016), "Long metal cable scattering features in space borne high resolution SAR images.", In Geoscience and Remote Sensing Symposium (IGARSS), IEEE International (pp.5982-5985), IEEE.

Civil Aviation Authority (CAA), (1999), "Pilot Intervention Times in Helicopter Emergencies.", CAA Paper No. 99001, U.K.

Cortes, C., Vapnik, V., (1995), "Support-vector networks.", Machine learning, 20(3), 273-297.

49

Csurka, G., Dance, C., Fan, L., Willamowski, J., Bray, C. (2004), "Visual categorization with bags of keypoints.", In Workshop on statistical learning in computer vision, ECCV (Vol. 1, No. 1-22, pp. 1-2).

Dalal, N., Triggs, B. (2005), "Histograms of oriented gradients for human detection.", In Computer Vision and Pattern Recognition, CVPR, IEEE Computer Society Conference on (Vol. 1, pp. 886-893), IEEE.

Donahue, J., Jia, Y., Vinyals, O., Hoffman, J., Zhang, N., Tzeng, E., Darrell, T. (2014), "Decaf : A deep convolutional activation feature for generic visual recognition.", In International conference on machine learning (pp. 647-655).

Duda, R. O., Hart, P. E., (1971), "Use of the Hough transformation to detect lines and curves in pictures.", (No. SRI-TN-36), SRI International Menlo Park CA, Artificial Intelligence Center.

Gao, Z., Wang, L., Zhou, L., Zhang, J., (2017), "HEp-2 cell image classification with deep convolutional neural networks.", IEEE journal of biomedical and health informatics, 21(2), 416-428.

He, K., Zhang, X., Ren, S., Sun, J., (2016), "Identity mappings in deep residual networks.", In European conference on computer vision (pp. 630-645), Springer, Cham.

Hu, F., Xia, G. S., Hu, J., Zhang, L., (2015), "Transferring deep convolutional neural networks for the scene classification of high-resolution remote sensing imagery.", Remote Sensing, 7(11), 14680-14707.

Ioffe, S., Szegedy, C., (2015), "Batch normalization : Accelerating deep network training by reducing internal covariate shift.", arXiv preprint arXiv : 1502.03167.

Krizhevsky, A., Sutskever, I., Hinton, G. E., (2012), "Imagenet classification with deep convolutional neural networks.", In Advances in neural information processing systems (pp. 1097-1105).

Li, Z., Liu, Y., Hayward, R., Zhang, J., Cai, J., (2008), "Knowledge-based powerline detection for UAV surveillance and inspection systems.", In Image and Vision Computing New Zealand, IVCNZ, 23rd International Conference (pp. 1-6), IEEE.

Li, Z., Liu, Y., Walker, R., Hayward, R., Zhang, J., (2010), "Towards automatic powerline detection for a UAV surveillance system using pulse coupled neural filter and an improved Hough transform.", Machine Vision and Applications, 21(5), 677-686.

Liu, Y., Mejias, L., Li, Z., (2012), "Fast powerline detection and localization using steerable filter for active UAV guidance.", International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences.

Luo, X., Zhang, J., Cao, X., Yan, P., Li, X., (2014), "Object-aware powerline detection using color and near-infrared images.", IEEE Transactions on Aerospace and Electronic Systems, 50(2), 1374-1389.

Ma, Q., Goshi, D. S., Shih, Y. C., Sun, M. T., (2011), "An algorithm for powerline detection and warning based on a millimeter-wave radar video.", IEEE Transactions on image processing, 20(12), 3534-3543.

Martinez, C., Sampedro, C., Chauhan, A., Campoy, P., (2014), "Towards autonomous detection and tracking of electric towers for aerial powerline inspection.", In Unmanned Aircraft Systems (ICUAS), International Conference on (pp. 284-295), IEEE.

Murphy, K. P. (2006), "Naive bayes classifiers.", University of British Columbia.

Ojala, T., Pietikainen, M., Maenpaa, T., (2002), "Multiresolution gray-scale and rotation invariant texture classification with local binary patterns.", IEEE Transactions on pattern analysis and machine intelligence, 24(7), 971-987.

Pan, C., Cao, X., Wu, D., (2016), "Powerline detection via background noise removal.", In Signal and Information Processing (GlobalSIP), IEEE Global Conference on (pp. 871-875), IEEE.

Penatti, O. A., Nogueira, K., dos Santos, J. A., (2015), "Do deep features generalize from everyday objects to remote sensing and aerial scenes domains?", In Proceedings of the IEEE conference on computer vision and pattern recognition workshops (pp. 44-51).

Sharif Razavian, A., Azizpour, H., Sullivan, J., Carlsson, S., (2014), "CNN features off the-shelf : an astounding baseline for recognition.", In Proceedings of the IEEE conference on computer vision and pattern recognition workshops (pp. 806-813).

Saito, H., Doshida, M., Takada, Y., Mine, Y., Hara, H., Kubomura, H.,& Yamamoto, J., (1992), "Obstacle warning lidar system using LD-pumped solid-state lasers.", In Conference on Lasers and Electro-Optics (p.CThI19), Optical Society of America.

Shan, H., Zhang, J., Cao, X., (2015), "Powerline detection using spatial contexts for low altitude environmental awareness.", In Integrated Communication, Navigation, and Surveillance Conference (ICNS), (pp. W2-1), IEEE.

Sarabandi, K., Pierce, L., Oh, Y., Ulaby, F. T., (1994), "Powerlines : Radar measurements and detection algorithm for polarimetric SAR images.", IEEE Transactions on aerospace and electronic systems, 30(2), 632-643.

Sarabandi, K., Park, M., (1999), "Millimeter-wave radar phenomenology of powerlines and a polarimetric detection algorithm.", IEEE Transactions on Antennas and Propagation, 47(12), 1807-1813.

Sarabandi, K., Park, M., (2000), "Extraction of powerline maps from millimeter-wave polarimetric SAR images.", IEEE Transactions on Antennas and Propagation, 48(12), 1802-1809.

Song, B., Li, X., (2014), "Powerline detection from optical images.", Neurocomputing, 129, 350-361.

Tian, F., Wang, Y., Zhu, L., (2015), "Powerline recognition and tracking method for uavs inspection.", In Information and Automation, IEEE International Conference on (pp. 2136-2141), IEEE.

Vengalattore T. N., Inderjit C., (2008), "Safety study of wire strike devices installed civil and military helicopters.", Department of Transportation Federal Aviation Administration Air Traffic Organization Operations Planning Office of Aviation Research and Development.

Yamaguchi, H., Kajiwara, A., Hayashi, S., (2000), "Power transmission line detection using an azimuth angular profile matching scheme.", In Radar Conference, The Record of the IEEE 2000 International (pp. 787-792), IEEE.

Yan, G., Li, C., Zhou, G., Zhang, W., Li, X., (2007), "Automatic extraction of powerlines from aerial images.", IEEE Geoscience and Remote Sensing Letters, 4(3), 387-391.

Yetgin, Ö. E., Gerek, Ö. N. (2013), "Cable and wire detection system for aircrafts.", In Signal Processing and Communications Applications Conference (SIU), 21st (pp. 1-4), IEEE.

Yetgin, Ö. E., Şentürk, Z., Gerek, Ö. N., (2015), "A comparison of line detection methods for powerline avoidance in aircrafts.", In Electrical and Electronics Engineering (ELECO), 9th International Conference on (pp. 241-245), IEEE.

Yetgin, Ö. E., Gerek, Ö. N., (2017), "PLD: Powerline detection system for aircrafts.", In Artificial Intelligence and Data Processing Symposium (IDAP), International (pp. 1-5), IEEE.

Yetgin, Ö. E., Gerek, Ö. N., (2018), "Automatic recognition of scenes with powerline wires in real life aerial images using DCT-based features.", Digital Signal Processing, 77, 102-119.

Yetgin, Ö.E., Gerek, Ö.N., (2017), "Powerline Image Dataset Extra, (Infrared-IR and Visible Light-VL) - Classified (Easy and Hard).", Mendeley Data.

Yetgin, Ö.E, Gerek, Ö.N., (2017), "Ground Truth of Powerline Dataset (Infrared-IR and Visible Light-VL).", Mendeley Data.

Yetgin, Ö.E., Gerek., Ö.N., (2017), "Powerline Image Dataset (Infrared-IR and Visible Light-VL).", Mendeley Data, v6.

Zeiler, M.D., Fergus, R., (2014), "Visualizing and understanding convolutional networks.", In European conference on computer vision (pp. 818-833), Springer, Cham.

Zhang, J., Shan, H., Cao, X., Yan, P., Li, X., (2014), "Pylon line spatial correlation assisted transmission line detection.", IEEE Transactions on Aerospace and Electronic Systems, 50(4), 2890-2905.

Wei, X., Zhang, R., Deng, B. (2008), "A recognition algorithm for high voltage transmission lines at horizontal polarization millimeter-wave radar.", In Information and Automation, ICIA, International Conference on (pp. 1172-1175), IEEE.

CURRICULUM VITAE

Surname, Name : Ömer Emre YETGİN

Language : English

Birthplace and Year : Istanbul / 1982

E-Posta : omeremreyetkin@gmail.com

Education and Professional Background :

- **Year, University, Institute / Faculty, Department, Department / Division**

  1) 2006, Erciyes University, Engineering Faculty, Department of Electronics Engineering,

  2) 2012, Gazi University, Technology Faculty, Department of Electronics and Computer Education,

  3) 2018, Eskisehir Technical University, Department of Electrical and Electronics Engineering.

- **Year, Duty, Institution, Unit**

  1) 2017, Captain, Turkish Armed Forces, Turkish Military Academy.

  2) 2018, Procurement Manager, Milmast A.S.

Publications and / or Scientific / Artistic Activities :

- **Year, Publication / Activity Type, Name and Information, Location**

  1. 2013, Conference Paper, Yetgin, O.E and Gerek, O,N., Cable and wire detection system for aircrafts, Signal Processing and Communications Applications Conference (SIU).

  2. 2014, Conference Paper, Yetgin, O.E and Gerek, O,N., Voiced-unvoiced classification of speech using autocorrelation matrix, Signal Processing and Communications Applications Conference (SIU).

  3. 2015, Conference Paper, Yetgin, O.E and Gerek, O,N., Segmentation performance comparison over HDR images, Signal Processing and Communications Applications Conference (SIU).

ner and Saliency Detection Methods for Powerline Detection, International Artificial Intelligence and Data Processing (IDAP).

5. 2017, Conference Paper, Yetgin, O.E., Gerek and O.N., PLD : Powerline Detection for Aircrafts, International Artificial Intelligence and Data Processing (IDAP).

6. 2017, Dataset, Yetgin, Ömer Emre and GEREK, Ömer Nezih, Powerline Image Dataset (Infrared-IR and Visible Light-VL), Mendeley Data, v7.

7. 2017, Dataset, Yetgin, Ömer Emre and Gerek, Ömer Nezih, Ground Truth of Powerline Dataset (Infrared-IR and Visible Light-VL), Mendeley Data, v8.

8. 2017, Conference Paper, Yetgin, O.E and Gerek, O,N., A comparison of corner and saliency detection methods for powerline detection, Artificial Intelligence and Data Processing Symposium (IDAP).

9. 2017, Journal Paper, Yetgin, O.E. and Gerek, O.N., Feature Extraction, Selection and Classification Code for Powerline Scene Recognition, DSP-SoftwareX Special Issue for Reproducible Research.

10. 2017, Journal Paper, Yetgin, O.E. and Gerek, O.N., Automatic Recognition of Scenes with Powerline Wires in Real Life Aerial Images Using DCT-Based Features, Journal of Digital Signal Processing (DSP).

11. 2018, Dataset, Yetgin, Ömer Emre and Gerek, Ömer Nezih, Powerline Image Dataset Extra (Infrared-IR and Visible Light-VL), Mendeley Data.

12. 2018, Yetgin. Ömer Emre, Benligiray, Burak and Gerek, Ömer Nezih, "Powerline Recognition from Aerial Images with Deep Learning", IEEE Transactions on Aerospace and Electronic Systems (TAES).

Awards :

- **Year, Degree, Award Name, Place**

1. 2012, BAP Project, 40.000 TL, Anadolu University.

2. 2016, TUBITAK, International Ph.D. Thesis Research Scholarship, 18.000 Euro.