DETERMINING RHEUMATOID ARTHRITIS AND OSTEOARTHRITIS DISEASES WITH PLAIN HAND X-RAYS USING CONVOLUTIONAL NEURAL NETWORK

A THESIS SUBMITTED TO
THE GRADUATE SCHOOL OF NATURAL AND APPLIED SCIENCES OF
ÇANKAYA UNIVERSITY

BY
KEMAL ÜRETEN

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF
MASTER OF SCIENCE
IN
COMPUTER ENGINEERING
DEPARTMENT

AUGUST 2019

Title of the Thesis: **Determining Rheumatoid Arthritis and Osteoarthritis Diseases with Plain Hand X-Rays Using Convolutional Neural Network**

Submitted by **Kemal ÜRETEN**

Approval of the Graduate School of Natural and Applied Sciences, Çankaya University.

Prof.Dr. Can ÇOĞUN
Director

I certify that this thesis satisfies all the requirements as a thesis for the degree of Master of Science.

Prof.Dr. Sıtkı Kemal İDER
Head of Department

This is to certify that we have read this thesis and that in our opinion it is fully adequate, in scope and quality, as a thesis for the degree of Master of Science.

Assoc. Prof.Dr. H. Hakan MARAŞ
Supervisor

Examination Date: 08.08.2019

Examining Committee Members

Committee Members

Assoc.Prof. Dr. H. Hakan MARAŞ (Çankaya Univ.)
Assoc. Prof. Dr. Nergiz ÇAĞILTAY (Atılım Univ.)
Asst. Prof. Dr. Abdül Kadir GÖRÜR (Çankaya Univ.)

STATEMENT OF NON-PLAGIARISM PAGE

I hereby declare that all information in this document has been obtained and presented in accordance with academic rules and ethical conduct. I also declare that, as required by these rules and conduct, I have fully cited and referenced all material and results that are not original to this work.

Name, Last Name : Kemal ÜRETEN

Signature :

Date : 21.08.2019

ABSTRACT

DETERMINING RHEUMATOID ARTHRITIS AND OSTEOARTHRITIS DISEASES WITH PLAIN HAND X-RAYS USING CONVOLUTIONAL NEURAL NETWORK

ÜRETEN, Kemal

M.Sc., Department of Computer Engineering

Supervisor: Assoc. Prof. Dr. Hadi Hakan MARAŞ

August 2019, 61 pages

Recent advances in computer technology have facilitated the acquisition of high-resolution images and processing of images. Convolutional neural network (CNN) is a branch of deep learning. CNN was first introduced in 1995 by LeCun, and in 2012, AlexNet won the ImageNet Large-Scale Visual Recognition Challenge (ILSVRC), after which there was rapid growth in deep learning applications. There are many successful studies using CNN especially in dermatology, pathology, radiology and ophthalmology. CNN highly successful in feature extraction and classification and requires less pre-processing. But in the CNN method, overfitting is an important problem that needs to be addressed and requires a large data set for training.

If there is not enough data for CNN training from scratch, previously trained CNN network from the natural image data set are used for transfer learning. Transfer learning is the use of a pre-trained model for a new problem. In recent years, there have been a few studies showing that CNN models trained with natural images have achieved successful results in the medical field.

Rheumatoid arthritis (RA) and hand osteoarthritis (OA) are two different diseases that cause pain, swelling, tenderness, loss of function in hand joints. In these diseases, affected joints and radiologic lesions show some differences. Treatment of

both diseases is also different. Conventional plain hand X-Rays (CR) are often used to diagnosis, differential diagnosis of RA and OA. The aim of this study is to develop a software that will help physicians for differential diagnosis of RA and OA from CR.

To the best our knowledge, this is the first study to distinguish between normal, hand OA and RA using plain hand radiographs. The efficiency of the created models was evaluated by using performance metrics such as accuracy, sensitivity, specificity and precision. In this study, pre-trained GoogLeNet, ResNet50 and VGG16 networks were used, transfer learning was applied. Successful results were obtained from all three pre-trained networks.

In this study, data augmentation, droupout, fine tuning, learning rate decay was applied to prevent overfitting. And during the training, no signs of overfitting were observed in the training chart.

Keywords: Rheumatoid Arthritis, Osteoarthritis, Convolutional Neural Network, Transfer Learning, Computer Aided Diagnosis

ÖZ

DÜZ EL RADYOGRAFİLERİ KULLANILARAK ROMATOİD ARTRİT VE OSTEOARTRİT HASTALIKLARININ KONVOLUSYONAL SİNİR AĞLARI İLE TANISI

ÜRETEN, Kemal

Yükseklisans, Bilgisayar Mühendisliği Anabilim Dalı

Tez Yöneticisi: Doç.Dr. Hadi Hakan MARAŞ

Ağustos 2019, 61 sayfa

Bilgisayar teknolojisindeki son gelişmeler, yüksek çözünürlüklü görüntülerin elde edilmesini ve görüntülerin işlenmesini kolaylaştırmıştır. Konvolüsyonel sinir ağı (CNN) derin öğrenmenin bir dalıdır. CNN ilk olarak 1995 yılında LeCun tarafından tanıtıldı ve 2012'de AlexNet, ImageNet Büyük Ölçekli Görsel Tanıma Mücadelesini (ILSVRC) kazandı ve ardından derin öğrenme uygulamalarında hızlı bir büyüme oldu. Özellikle dermatoloji, patoloji, radyoloji ve oftalmoloji alanlarında CNN kullanımı ile ilgili birçok başarılı çalışma vardır. CNN özellik çıkarma ve sınıflandırma konusunda oldukça başarılı ve daha az ön işleme ihtiyaç duyar. Ancak CNN yönteminde, overfitting ele alınması gereken önemli bir sorundur ve CNN eğitimi için çok sayıda veri gerekir.

Sıfırdan CNN eğitimi için yeterli veri olmadığında, daha önce doğal görüntü veri setinden eğitilmiş CNN ağı öğrenme transferi için kullanılır. Öğrenme transferi, yeni bir problem için önceden eğitilmiş bir modelin kullanılmasıdır. Son yıllarda, doğal görüntüler ile eğitilmiş CNN modelleri ile tıp alanında başarılı sonuçlar elde edildiğini gösteren birkaç çalışma vardır.

Romatoid artrit (RA) ve el osteoartriti (OA), el eklemlerinde ağrı, şişme, hassasiyet, fonksiyon kaybına neden olan iki farklı hastalıktır. Bu hastalıklarda, etkilenen eklem

ve radyolojik lezyonlar bazı farklılıklar göstermektedir. Her iki hastalığın tedavisi de farklıdır. Geleneksel düz el radyografileri (CR) RA ve OA ayırıcı tanısı için sıklıkla kullanılır. Bu çalışmanın amacı, hekimlere CR kullanılarak RA ve OA'nın ayırıcı tanısında yardımcı olacak bir yazılım geliştirmektir.

Bilgimize göre, bu çalışma düz el radyografileri kullanarak normal, el OA ve RA'yı ayırt eden ilk çalışmadır. Oluşturulan modellerin verimliliği, doğruluk, hassasiyet, özgüllük ve kesinlik gibi performans ölçütleri kullanılarak değerlendirildi. Bu çalışmada, önceden eğitilmiş GoogLeNet, ResNet50 ve VGG16 ağları kullanılmış, öğrenme transferi uygulanmıştır. Önceden eğitilmiş her üç ağdan da başarılı sonuçlar alındı.

Bu çalışmada, overfittingi önlemek için veri büyütme, droupout, ince ayar, öğrenme hızı azalması uygulanmış ve eğitim çizelgesinde overfitting izleri gözlenmemiştir.

Anahtar kelimeler: Romatoid Artrit, Osteoartrit, Convolutional Sinir Ağları, Öğrenme Transferi, Bilgisayar Destekli Tanı

ACKNOWLEDGEMENTS

The author wishes to express his deepest gratitude to all the faculty members of Çankaya University Computer Engineering for giving the opportunity to graduate in computer engineering.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF SYMBOLS/ABBREVIATIONS

RA **:** Rheumatoid arthritis

OA **:** Osteoarthritis

CNN **:** Convolutional Neural Network

CR **:** Conventional plain hand radiographs

ReLU **:** Rectified linear unit

sgdm **:** stochastic gradient descent with momentum

CAD **:** Computer Aided Diagnosis

MRI **:** Magnetic Resonance Imaging

FC **:** Fully connected

DIP **:** Distal interphalangeal

PIP **:** Proximal interphalangeal

MCP **:** Metacarpophalangeal

ESR **:** Erythrocyte sedimentation rate

CRP **:** C-reactive protein

RF **:** Rheumatoid factor

Anti-CCP **:** Anti-Cyclic citrullinated protein

TNF : Tumor necrosis factor

CHAPTER 1

INTRODUCTION

1.1 Background of the Study

Medical imaging is a fundamental method in the diagnosis, early diagnosis and differential diagnosis of diseases. Computer Aided Diagnosis (CAD) methods are used to assist physicians in the interpretation of medical images (direct X-rays, Magnetic Resonance Imaging images, digital pathology images). In CAD method, the physician first makes a routine evaluation on the image, re-evaluates own interpretation with the help of the CAD system and makes the final decision. Thus the doctor receives a second objective opinion aid. Some research suggests that the inclusion of the CAD system in the diagnostic process provides quantitative support for clinical decisions by reducing inter-observer variations.

Convolutional neural networks (CNN) are deep learning architectures and firstly introduced by LeCun et al [1]. CNN application is a state-of-the-art in image processing. CNN highly successful in feature extraction and classification, and requires less pre-processing. However, CNN requires a large dataset for training. Therefore it is a challenge to train a deep CNN from scratch (or full training). Millions of images are required for a successful CNN application [2]. If there is not enough data available, data augmentation and transfer learning are applied [3], [4], [5].

Using CNN application, diagnosis of brain tumors from MR images, diagnosis of interstitial lung disease using CT images, and diagnosis of colon cancer using histopathological images are only a few of these successful studies. Nowadays, there are many different pre-trained networks that are used successfully for transfer learning developed using ImageNet database. Some of those are AlexNet [6], VGG-16 [7], GoogLeNet [8], SqueezeNet [9], ResNet [10], inceptionv4 [11]. Transfer learning is the use of a previously trained model on a new problem. It is now very popular method in the field of Deep Learning. Finding sufficient data in the medical field is particularly difficult, as some diseases are rare or sufficient data cannot be

obtained for reasons such as patient confidentiality, so there is not currently sufficient dataset available in many medical subjects. In this case, previously trained CNN models from the natural image dataset are used to apply transfer learning [5].

Rheumatoid arthritis (RA) and hand osteoarthritis (OA) are two different diseases that cause pain, swelling, tenderness, and loss of function in hand joints. When the joints affected in these diseases, the characteristics of the resulting lesions show some differences. Treatment of both diseases is also different. In order to diagnosis, differential diagnosis and monitor the activities of these diseases, plain hand radiographs are frequently used imaging methods. Plain X-Rays are relatively inexpensive and easily accessible methods of examination. Radiological changes occur relatively late in these diseases. Evaluation of hand X-Rays requires experience. In addition, some findings may be overlooked due to reasons such as work intensity, fatigue, carelessness, and insufficient time. In some centers even plain radiography is available; Radiologist, Rheumatologist, or Physical Therapy Specialist may be absent.

If RA and OA can be differentiated by hand radiography, the patient with RA is referred to the specialist for early treatment. With early RA treatment, permanent deformities seen in RA can be prevented. And with early diagnosis of OA, the patient's quality of life and improvement in the course of the disease are improved by simple procedures such as patient education and lifestyle modification. Thus, general practitioners working in primary health care institutions may take an active role in OA treatment. In addition, experienced physicians, such as radiologists and rheumatologists, may also use CAD methods to make their final decisions.

1.2 Objectives

The aim of this study is to develop a CAD system to assist physicians in the differentiation of patient hand radiography with rheumatoid arthritis and patient hand radiography with osteoarthritis, and normal hand radiography. Thus, physicians will get a second objective aid when evaluating hand radiographs. This CAD system may be used by experienced doctors such as radiologists and rheumatologists as well as general practitioners.

1.3 Organization of the Thesis

Chapter 1 contains the introduction part and aims of this thesis,

Chapter 2 includes information about rheumatoid arthritis, osteoarthritis diseases, convolutional neural networks, computer-aided diagnosis. In addition, a brief summary of previous studies with medical images using neural networks methods is presented,

Chapter 3 provides information about data acquisition, classification, preprocessing, number of data used for testing and training, and the study design,

Chapter 4 presents the results,

Chapter 5 includes discussion and conclusion part.

CHAPTER 2

OVERVIEW OF STUDY TOPICS

This section provides information about rheumatoid arthritis and osteoarthritis diseases, convolutional neural networks and computer aided diagnosis. In the medical field, a brief summary of previous studies with neural network methods is presented.

2.1 Rheumatoid Arthritis

Rheumatoid Arthritis (RA) is the most common inflammatory arthritis, seen in 0.5-1% of the population, with an annual prevalence of 40 / 100,000. The average age of onset is between 30 and 50 years and is more common in women than in men. It can affect all synovial joints, and especially the metacarpophalangeal (MCP), proximal interphalangeal (PIP) and wrist joints in hand [12]. RA is a systemic inflammatory connective tissue disease, and it can cause different complaints in the skin, eye, lung, heart and neurological system besides joints. The diagnosis of RA is made by the evaluation of the patient's examination findings, laboratory results and imaging findings of the patient together. Plain hand radiographs are the first line and most commonly used imaging methods for diagnosis, differential diagnosis of RA and monitoring of disease activity. Plain radiographs are relatively inexpensive and easily accessible imaging methods.

2.1.1 Etiopathogenesis

In the etiology of rheumatoid arthritis, environmental factors, hormonal, genetic, infectious causes are accused. In genetically susceptible individuals, more than one factor may induce polyarthritis. Once started, the process is self-sustaining. Various bacterial and viral infections are suspected in the onset of the disease, but to date an infectious agent has not been established. RA occurs as a result of complex interaction between genetic and environmental factors, and abnormal activation of the immune response results in synovial hyperplasia and bone destruction [13].

In the pathogenesis of RA, T cells, B cells, synovial cells, fibroblasts and macrophages play an important role, synovial tissue is the main target of inflammation and eventually synovitis develops. Once the autoimmune process begins, the synovium in RA is organized into an invasive tissue capable of degrading cartilage and bone. Rheumatoid synovium (it is called 'pannus') shows many features of a locally invasive tumor. Cytokines play an important role in the pathogenesis of RA. Cytokines are proteins that enable chemical communication between cells, growth and differentiation of cells, and regulation of immune response. They are secreted by cells of the immune system. Cytokines that play major role in RA pathogenesis are interleukin-1 (IL-1) and tumor necrozis factor-α (TNF-α) [14].

2.1.2 Clinical findings

RA usually starts insidiously, causing pain, stiffness and swelling in the joints. RA may involve all synovial joints, leading to symmetrical involvement. RA primarily involves the metacarpophalangeal (MCP) and proximal interphalangeal (PIP) joints, and the wrist joints. In RA, other synovial joints of the upper and lower extremities, such as metatarsophalangeal (MTP) joints of the toes, ankles, elbows, shoulders and knees, are also frequently affected. Morning stiffness lasting more than 1 hour is an important finding in diagnosis [15], [16], [17]. In the early stages of the disease, there is symmetrical, fusiform soft tissue swelling and widening of the joint space due to synovial inflammation in the joints; followed by periarticular osteoporosis, concentric joint space narrowing and marginal erosions, central erosions and fibrous ankyloses. If not treated adequately, later stages of the disease may develop intense marginal and subchondral erosive changes, loss of joint space, bone ankyloses and deformities due to cartilage and bone erosion. As a result, deformities such as ulnar deviation, swan neck deformity, boutonniere deformity develop [18], [19], [20]. (Figure 1 shows hand pictures of 2 patients with rheumatoid arthritis)

Figure 1 Hand pictures of 2 patients with rheumatoid arthritis

2.1.3 Laboratory findings

Acute phase reactants such as erythrocyte sedimentation rate (ESR) and C-reactive protein (CRP) may be high in patients with RA. Patients with RA have anemia and thrombocytosis, sometimes mild lococytosis, consistent with chronic inflammation. There may be leukopenia and thrombocytopenia due to Felty syndrome or drugs. Liver and kidney function tests are usually normal. Rheumatoid factor (RF) and Anti-Cyclic citrullinated protein (anti-CCP) autoantibodies are positive in 70% of patients with RA. The presence of RF, autoantibodies that react with the Fc portion of IgG, is not specific for RA. RF may be positive in 5-10% of healthy individuals. In addition, RF may be positive during the course of some connective tissue diseases such as Systemic Lupus Erythematosus (SLE), Sjogren's syndrome, and some infectious diseases such as Hepatitis B, Tuberculosis, and Leprosy. Anti-CCP antibodies in RA are a new and highly specific marker. The specificity of anti-CCP autoantibodies for the diagnosis of RA varies between 92% and 96%, and its sensitivity varies between 53-68%. However, anti-CCP antibodies may also be positive in some diseases other than RA such as SLE, Sjogren, Juvenile Idiopathic arthritis [20].

2.1.4 Imaging in rheumatoid arthritis

The imaging methods commonly used for the diagnosis of RA are plain radiographs, magnetic resonance imaging (MRI) and ultrasonography. Plain hand radiographs are the most commonly used imaging modality in RA diagnosis, differential diagnosis and follow-up. Plain radiographs are relatively inexpensive and easily accessible imaging methods. However, radiographic changes occur relatively late in rheumatoid

arthritis [19]. In the early stages of disease, symmetrical, fusiform soft tissue swelling and joint space widening is seen, due to synovial inflammation; following this, periarticular osteoporosis, concentric joint space narrowing, and marginal erosions, central erosions and fibrose ankyloses may develop. Cortical erosion indicates bone loss and starts at bare areas where the joint corners are free of cartilage. In early RA, erosions may occur around the ulnar styloid and scaphoid bone. In advanced RA cases with cartilage loss, subluxation and dislocation of the joints may occur. If not treated adequately, later stages of the disease may involve intense marginal and subchondral erosive changes, loss of joint space, bony ankyloses and deformities [18], [20], [21], [22]. (Figure 2 shows plain hand radiographs of 2 patients with rheumatoid arthritis)

Figure 2 Hand radiographic image of 2 patients with rheumatoid arthritis

To determine the damage caused by RA in the hand and wrist can be used Modified Sharp score, which is assessed in 18 points and 17 joints in the hand and wrist. Bone erosion is assessed with a score of 0 to 5, and joint space narrowing is assessed with a score of 0 to 4. Higher score means more damage [23].

Magnetic Resonance Imaging (MRI) is a medical technique used to detect and identify differences between healthy and diseased tissues using radio waves within the generated strong magnetic field. MRI can evaluate all structures involved in RA, including synovium, effusion, cartilage, bone, ligaments, tendons, and tendon sheath, and may help to investigate patients with atypical presentation or early disease. Bone erosions, tenosynovitis, synovial proliferation can be detected early by MRI; it is a reliable, radiation-free, but expensive imaging method [24], [25].

Ultrasonography is a diagnostic method obtained by using high frequency sound waves. The sound waves sent from the device are detected by the same device after

being reflected from the patient's body and the healthy and diseased tissues are distinguished by reflection differences. Ultrasonography is a widely available, relatively inexpensive imaging method and the patient is not exposed to ionizing radiation. Ultrasonography can detect synovial thickening, joint effusion and superficial erosions. Ultrasonography also works to distinguish between inflammatory and non-inflammatory arthritis. However, it is a time-consuming and operator-dependent imaging method. (Table 1 shows advantages and disadvantages of radiographic methods used for assessing RA) [19], [24], [25].

Table 1 Advantages and Disadvantages of Radiographic Methods

Radiographic methods	Advantages	Disadvantages
Conventional radiography	- Inexpensive - Easy accessible - Standardization available - Easy reproducibility - Valid assessment method	- Ionizing radiation - Insensitive to early bone damage - Inadequate evaluation of soft tissues
Magnetic resonance imaging	- Safe method - No ionizing radiation - High sensitivity - Assessment of all affected structures - Determination of bone edema	- Expensive method - Limited availability - MRI exams require extended periods of time
Ultrasonography	- Non-invasive method - Relatively inexpensive - No ionizing radiation - Detection of inflammatory changes	- Operator-dependent - Time-consuming - No standardization

2.1.5 Diagnosis

RA is diagnosed by evaluating the patient's history, physical examination findings, laboratory results and imaging findings together. The American Rheumatism Association (ARA) 1987 RA classification criteria and the American College of Rheumatology / European League Against Rheumatism Collaborative Initiative (ACR / EULAR) 2010 RA classification criteria are used for diagnosis of RA [26], [27].

2.1.6 Differential diagnosis

The differential diagnosis of RA is made by the following diseases; Osteoarthritis, viral polyarthritis, systemic rheumatic diseases (SLE, Sjogren syndrome, myositis, sarcoidosis), palendromic rheumatism, hypermobility syndrome and fibromyalgia, reactive arthritis, Lyme arthritis, Psoriatic arthritis, Polymyalgia rheumatica, crystal arthritis.

2.1.7 Treatment

Early diagnosis is important, early use of disease-modifying antirheumatic drugs (DMARD) may provide remission. The patient should be informed about the disease and psychosocial support should be provided in all stages of the disease. Drug treatment is started with DMARDs (methotrexate, sulfasalazine, hydroxychloroquine, leflunomide, corticosteroids). AntiTNF drugs (infliximab, etanercept, adalimumab, golimumab), B-cell therapies (rituximab), T-cell therapies (abatacept), tofacitinib are used if adequate response is not obtained with DMARDs [28], [29].

2.2 Osteoarthritis

Osteoarthritis (OA) is the most common degenerative arthritis and affects 10% of men and 13% of women. It is rare before the age of 40. With aging, changes in the articular cartilage occur, resulting in reduced endurance. Therefore, the incidence of OA increases with age. In Western societies, the prevalence of OA is expected to increase due to the aging population. OA leads to decrease in quality of life and functional losses in the advanced stages. OA may develop in a particular joint due to various stress factors, risk factors and genetic factors. OA causes osteophytes, subchondral sclerosis and asymmetric joint space narrowing (JSN) in hand joints [30], [31].

2.2.1 Etiopathogenesis

There are some risk factors for the development of OA, such as age, gender, overweight / obesity, joint trauma / sports injuries (associated joint instability and muscle laxity), certain occupations that cause recurrent stress on a particular joint, genetics, bone deformities, metabolic diseases (eg diabetes), endocrine disorders and other rheumatic diseases such as RA and gout. The risk of OA increases with aging, and is more common in women. Some of the risk factors are irreversible risk factors for the development of OA. The risk of OA development can be prevented by obesity and avoidance of sedentary life [32]. Repeated use of the joint while working increases the risk of OA. In particular for hand OA, occupations that require repetitive movement of the hand, such as labor job and clothing industry workers are at higher risk [33], [34]. OA is classified as primary and secondary OA. If osteoarthritis develops without a history of previous trauma or predisposing cause, it is called primary OA. On the other hand, if OA develops in the joint affected by

trauma or due to a pre-existing condition / disease, it is called secondary OA [34]. OA is a biomechanical disease but also has inflammatory and metabolic components. Although all joint tissues play a role in the onset and progression of OA, joint cartilage is the most important. The main function of articular cartilage is to provide a smooth and oiled surface for articulation and to facilitate the transmission of loads with low friction. Articular cartilage is avascular, non-innervative tissue and has very low cell. Cartilage is fed through the synovium through diffusion. Cartilage contains chondrocytes and collagen matrix, and 65-80% of its volume is water. Production and destruction of normal cartilage are in balance. Disruption of the balance leads to the change of cartilage structure, cartilage loss occurs; hence the joint space is reduced. The result is bone damage, subchondral sclerosis, and new bone formation called osteophytes [35]. Various inflammatory cytokines such as interleukin-1 beta (IL-1β) and tumor necrosis factor alpha (TNF-alpha) and mediators (metalloproteinases) produced by joint tissue are involved in the pathogenesis of osteoarthritis [32].

2.2.2 Clinical findings

OA often affects the knees, hips and hand joints. In particular, hand OA affects distal interphalangeal (DIP) joints, the first carpometacarpal joints, proximal interphalangeal (PIP) joints, less commonly the second and third metacarpophalangeal (MCP) joints in the hands [36]. Clinically, it is characterized by joint pain, tenderness, crepitus, bony swelling, and instability in the joint, stiffness and limitation of movement. Joint pain increases in the afternoon and is often associated with activity, and persistent pain may be seen in later stages of the disease. In severe OA, night pain may be seen disrupting sleep [37]. Joint stiffness, which lasts less than half an hour, is especially common in the morning. In the later stages of the disease, a sense of instability in the joint is common, as well as "cracking" that can be heard and palpable on a joint during active or passive movement of the joint. This may be caused by effusion, capsular contractures, muscle spasm or weakness, loose joints, mechanical constraints, and joint deformity [34]. (Figure 3 shows hand pictures of 2 patients with osteoarthritis)

Figure 3 Hand pictures of 2 patients with osteoarthritis

2.2.3 Imaging

In hand OA, osteophytes occur on the joint corners. Enlargement on DIP joints is called Heberden nodules, and enlargement on PIP joints are called Bouchard nodules, the first carpometacarpal joint is also frequently involved. Imaging methods used in the diagnosis of OA are x-rays, magnetic resonance imaging (MRI) and ultrasonography. Conventional plain hand radiographs (CR) are the gold standard diagnostic method for diagnosing hand OA. CRs are successful to diagnose and differentiate OA; it is also successful to demonstrate the severity of the disease, and to monitor the activity of the disease [38], [39]. CRs are relatively inexpensive, most commonly used and easily accessible imaging methods. Osteophytes, joint space narrowing (JSN), subchondral sclerosis, and pseudocystic areas, subchondral erosions, subluxations and loss of joint space can be detected with CR. Joint space loss in OA is non-uniform and asymmetric [40]. Kellgren-Lawrence (KL) scoring has been used radiologically for more than 50 years to evaluate OA findings and staging the severity of the disease. In Kellgren-Lawrence scoring, joint space narrowing, osteophytes and sclerosis are evaluated, and scores between 0-4 are given [34]. (Figure 4 shows plain hand radiographs of 2 patients with osteoarthritis)

Figure 4 Hand radiographic images of 2 patients with osteoarthritis

2.2.4 Diagnosis

The diagnosis of OA is based on clinical and radiological findings. Examination of hand OA can detect joint instability and ankylosis, Heberden nodes and Bouchard nodes and joint deformities. In OA, changes reflecting the acute phase reaction such as anemia, thrombocytosis, increased erythrocyte sedimentation rate, and significant immunological abnormalities are not detected. Imaging methods are used to make the diagnosis, to determine the severity of the disease and to exclude other pathologies. CRs, MRI and ultrasound are the imaging methods that can be used for this purpose [34], [41], [42]. OA is diagnosed according to the classification criteria proposed by the American College of Rheumatology [36].

2.2.5 Differential diagnosis

OA differential diagnosis is made with Rheumatoid arthritis, Psoriatic arthritis, Crystalline arthritis, Hemochromatosis, Infectious arthritis.

2.2.6 Treatment

Treatment of OA includes patient education, use of assistive devices such as splints, lifestyle changes, removal of predisposing factors, and, if necessary, a change in occupation. The load on the joint should be reduced and balanced. Various physiotherapy methods such as hot and cold application, low frequency currents, traction and massage can be applied. The patient is allowed to exercise (stretching exercises, strengthening exercises). Analgesics, topical and oral nonsteroidal anti-inflammatory drugs (NSAIDs) are used in the pharmacological treatment of OA, and disease modifying drug development studies are ongoing. If necessary, surgical methods such as osteotomy, arthroplasty, arthroscopic irrigation, debridement, and

arthrodesis can be applied to the joints. Early diagnosis improves the patient's quality of life and reduces the course of the disease [43], [44].

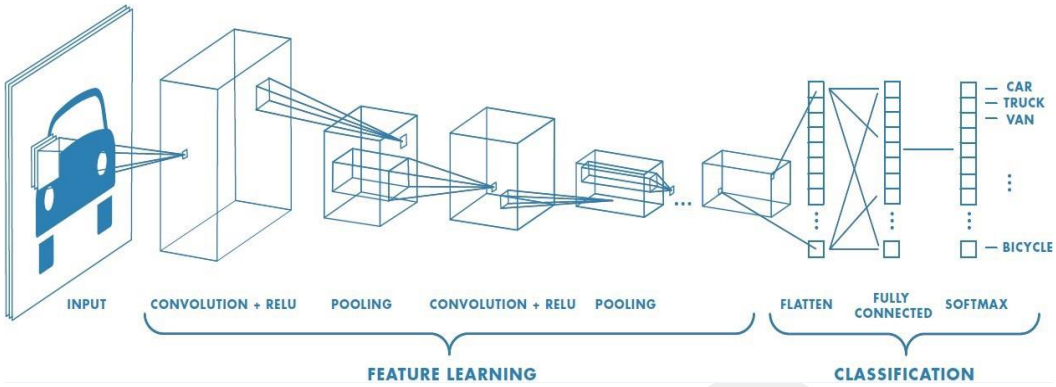
2.3 Convolutional neural networks

Deep learning is a sub-area of machine learning. Successful results have been obtained with deep learning methods in the area of medical image analysis. Deep learning methods does not require handcrafted feature extraction, this allows successful classification of medical images and diagnosis of diseases. Convolutional neural networks (CNN), Deep Belief Networks, Deep Boltzmann Machines are the most popular deep learning methods [45].

Convolutional neural networks (CNN) were first introduced by LeCun et al., and have many successful applications in image and video recognition, suggestion systems, image classification, medical image analysis and natural language processing. CNN is particularly successful in feature extraction and classification. CNNs are multilayer perceptron's designed to require minimal preprocessing [1], [46], [47]. This independence to preprocessing and human effort is a big advantage. However, a large amount of annotated training data is required for CNN training. The network learns directly from the data, we cannot influence which characteristics are learned [48].

CNNs were designed with inspiration from the biological structure of the animal visual cortex. There are simple and complex cells in the visual cortex. These cells were found to be active based on the lower regions of the visual field. These sub-regions are called receptive fields. Simple cells, for example, are activated when they define basic shapes in a fixed area and at an angle. Complex cells have larger receptive fields, and their outputs are not sensitive to the specific position in the area [49].

Neural networks have three layers: an input layer, a hidden layer and an output layer. The data are presented to the network via the input layer, the desired output is obtained from the output layer, and the associated rules are learned in the hidden layer [50]. The hidden layers of a CNN consist of convolution layers, pooling layers, fully connected layers and normalization layers [51]. (Figure 5 shows Convolution Neural Network architecture)



source : https://towardsdatascience.com/basics-of-the-classic-cnn-a3dce1225add

Figure 5 Architecture of Convolutional Neural Network

2.3.1 Convolutional layer

The core building block of the CNN is the convolutional layer. The role of a convolution layer is to detect local features in the input feature maps. A feature on the image can be computed by convolution. The Convolution layer has filters (Kernel). Convolution filters are applied to small areas of the input image. The input field that overlaps the filter is called the receptive field. At the end of the convolution process, an output called feature maps is obtained [52]. During the convolution process, the filters are shifted on the image from left to right and from top to bottom. During the convolution operation, the overlapping input and filter elements are multiplied at each step and then all elements are summed to produce a scalar output (Figure 6 illustrates the convolution operation). When this process is applied to the whole image, a new image is obtained. Multiple filters are applied on the image, and are obtained feature maps as much as the number of filters is obtained. When Stride 1, the filters are shifted one pixel at a time. When Stride is 2 (or rarely 3 or more), the filters shift 2 pixels on the image. Receptive fields overlap less and the resulting output volume has smaller spatial dimensions, but data loss is greater when stride is greater.

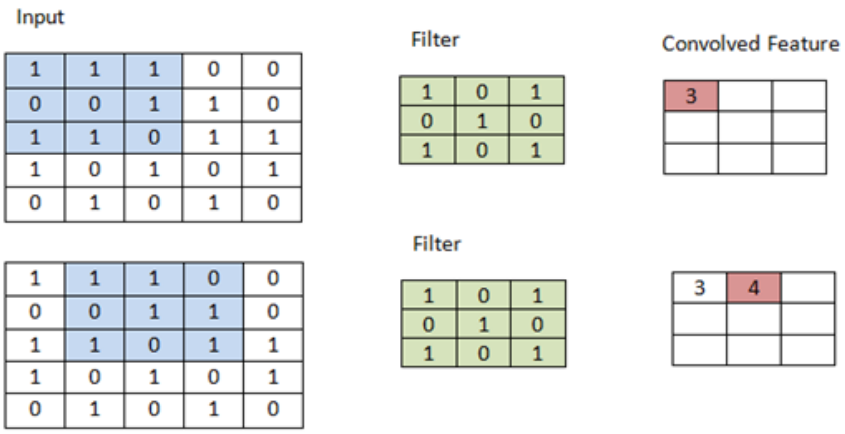


Figure 6 Convolution operation

In the convolution layer, the number of filters is variable, while the layers next to the input layer tend to have fewer filters, while the higher layers may have more filters. The convolution process results in a different number of feature maps, depending on the number of filters used.

Filter size is highly effective on learning. Deciding the size of the kernel determines how wide the data will affect each other. Generally 3x3, 5x5, 7x7 kernels are used. Using a large kernel will cause the image to be smaller after convolution is applied. As this causes loss of knowledge, a small kernel such as 3x3, 5x5 is generally used. Filters in the first layers of the network often define low-level features, such as edges and lines. Filters in the last convolution layers of the network identify more complex features such as shapes, texture and defined concepts [5].

Padding - After convolution operation, the output dimensions are smaller than the original input dimensions. Padding is used to make the output image the same size as the input. Padding provides control of the spatial size of the output volume. One way to padding is to add zeros outside the boundaries of the output image [53]. (Figure 7 illustrates the padding operation).

0	0	0	0	0	0
0	8	12	3	2	0
0	4	15	5	7	0
0	14	9	22	4	0
0	5	6	14	3	0
0	0	0	0	0	0

Figure 7 Padding operation

2.3.2 Pooling layers

In the CNN architecture, each layer of convolution is followed by a pooling layer. Pooling simplifies output by performing a non-linear sub-sampling, reducing the number of parameters the network needs to learn. It helps reduce operational complexity and reduces the size of the input. As a result, the pooling layer reduces the likelihood of overfitting the trained model [5], [51].

There are some non-linear functions to implement pooling; max pooling, average pooling, and L2-norm pooling. Maxpooling is the most commonly used. Usually 2x2 size filters are used; the largest element in the receptive field of the input is selected. Thus, it resizes spatially, depth dimension remains unchanged. Selecting larger filters (3x3, 4x4) significantly reduces the size of the image and may result in loss of knowledge. Average pooling uses the average value from each of a set of neurons in the previous layer. Max pooling works better in practice [52], [53]. (Figure 8 illustrates the max pooling operation)

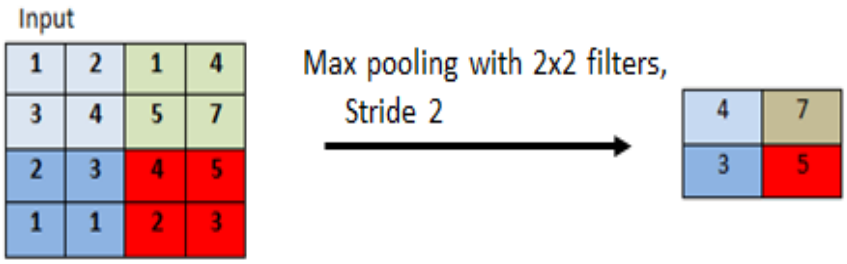


Figure 8 Max pooling operation

2.3.3 Fully connected layer

Finally, after several convolutions and pooling layers, fully connected (FC) layer is followed. FC layers are used as final layers in CNNs. The last layer outputs the estimations. Multi-class networks use Softmax classifier in the final layer while binary classification is done using the sigmoid function. Softmax function is a logistic activation function used for multiple classifications. Softmax layer must have the same number of nodes as the output layer. Softmax returns decimal probabilities for each category in the data set, that is, the value of the output sum of 1. The FC layer basically receives an input volume and returns an N-dimensional vector, which is the number of classes the program has to select. For example, in a digit classification program, N is 10 because there are 10 digits. Basically, an FC layer

looks at which high-level properties are most strongly associated with which class and obtain the correct possibilities [54].

2.3.4 ReLU Layer

Convolution operation is a linear process, so convolutional layers can only model linear dependence. The main purpose of nonlinear activation functions is to introduce nonlinearity in neural networks. After each convolution layer, a rectified linear unit (ReLU) is applied to model to give non-linearity. ReLU allows faster and more efficient training by mapping negative values to zero and maintaining positive values. It enhances the nonlinear properties of the decision function and the overall network without affecting the receptor areas of the convolution layer. Deep learning methods are more effective than other methods in solving non-linear problems [5], [55]. This layer implements function f (x) = max (0, x). The equation for the output of the function can be written as follows (1). X when X => 0, and 0 when X <0.

$$ReLU(x) = \begin{cases} 0 \; if \; x < 0 \\ x \; if \; x \geq 0 \end{cases} \quad (1)$$

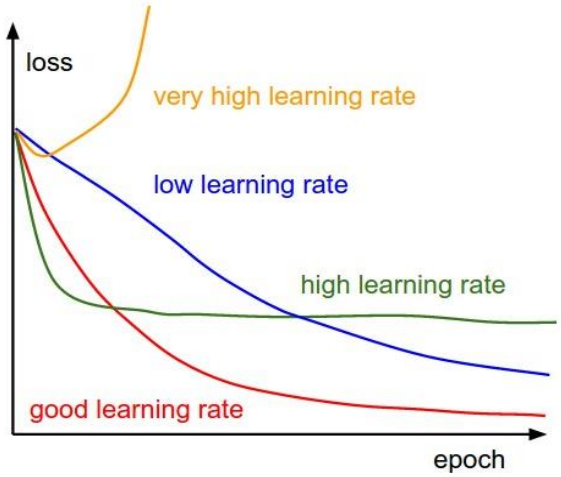
However, to be zero of all negative values reduces the ability of the model to be trained according to the data. A common problem is the dying ReLU, which means that the function will give zero for most inputs, so that the problem of missing gradients emerges. To prevent this, leaky ReLU is used. Other functions are also used to increase non-linearity, eg saturated hyperbolic tangent f (x) = | tanh (x) |, and the sigmoid function f (x) = 1 / (1 + e-x). ReLU is often preferred to other functions [5], [56].

2.3.5 Dropout

Another normalization technique is dropout. Dropout prevents overfitting and significantly increases training speed. Dropout helps to ensure that the entire setup does not depend on a particular neuron or connection by turning randomly selected neurons "on" and "off" during training, and thus the network is forced to learn new ways. It has been observed that dilution of nodes below a certain threshold in FC layers improves performance. In training stages, usually 20-50% of neurons are reduced; it must be low for input nodes (20%) because high dropout on input nodes leads to data loss. A very low dropout value has little effect and a very high value results in inadequate learning by the network [57].

2.3.6 Learning Rate

In deep learning, the parameters are updated by the backpropagation process. The learning rate defines how quickly a network updates its parameters. It is important to select the value of the learning rate as the step size directly affects the learning process. The higher learning rate, the larger steps updates in weight and therefore it may take less time for the model to converge to the optimum weight set. However, a very high learning rate may result in jumps that are too large and not precise enough to reach the optimal point. With a very small learning rate, training may take much longer, and may stay at the appropriate local optimum and never converge. The optimal solution for learning rate is to keep the learning rate high at the beginning and to decrease it gradually [58]. (Figure 9 shows the effects of different learning rates)



source : http://cs231n.github.io/neural-networks-3/

Figure 9 Effects of different learning rates

2.3.7 Overfitting

A lot of data is required for CNN training from scratch. Training with insufficient number of data can result with overfitting. Overfitting means that the network learns to identify each image with its own result, rather than the resulting model (memorising). When new images are introduced to the network, the model cannot interpret them correctly and accuracy decreases [59]. To control overfitting during training, the data is usually divided into training, validation and test sets. This data division is used to estimate the last generalization error. The test set is not displayed to the network until the model is trained. Different regularization techniques such as reducing the complexity of the model, dropout, weight loss and data increase are

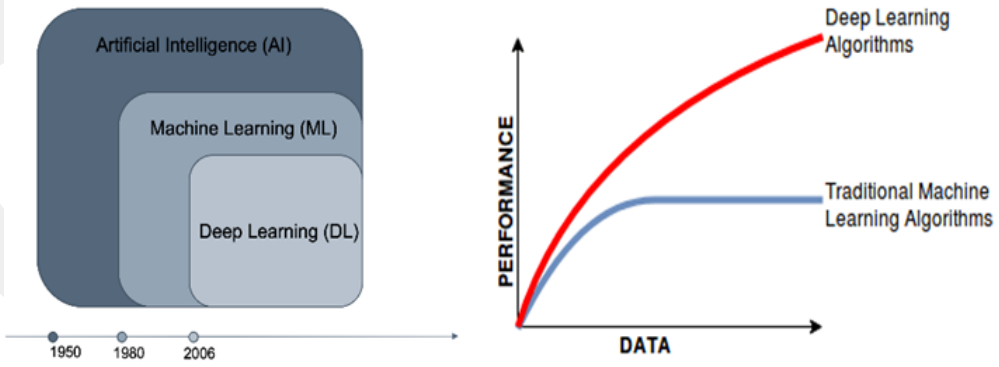
applied to avoid overfitting. One of the simplest ways to prevent overfitting is to stop training before the overfitting process takes place. Another simple way to avoid overfitting is typically to limit the number of hidden units on each layer or to limit the network depth. Thus, the number of parameters is limited; this reduces the complexity of the function that the data can perform, thus limiting the amount of overfitting [60].

2.3.8 Data Augmentation

The performance of the deep learning method depends largely on the number of data. When the dataset is small, the possibility of overfitting increases. If sufficient data is not available, the appropriate solution is data augmentation. Data augmentation is to generate more training data than an existing set of data. Various amplification techniques are used for data augmentation such as random cropping, color flicker, image flip, and random rotation [61].

2.3.9 Transfer learning

CNN highly successful in feature extraction and classification and requires less pre-processing. Disadvantages of CNN are overfitting, requires big dataset to training, and computational expensive. (Figure 10 shows a comparative graph of learning algorithms)



source : https://www.datacamp.com/community/tutorials/machine-deep-learning

Figure 10 Comparison of learning algorithms

For CNN training from scratch, a large amount of labeled training data is required. For a successful result, the quality of the images, enough data, and the appropriate CNN design are important [48], [51], [62]. When there is not enough data available,

data augmentation and transfer learning can be applied. Transfer learning is the use of a previously trained model on a new problem. Nowadays it is a state-of-the-art method in the field of deep learning. It allows training Deep Neural Networks with relatively little data.

This is very useful because most real-world problems typically do not have millions of labeled data sets to train such complex models. It is particularly difficult to find sufficient data in the medical field where specialist annotations are expensive and some diseases (egg, lesions) are scarce or sufficient data cannot be obtained for reasons such as patient confidentiality, so there is not currently sufficient dataset available in many medical fields. In this case, previously trained CNN models from a natural image data set or a different medical field are used for a new medical task [5]. ImageNet is a project that contains 1.4 million images in 1000 separate categories. Nowadays there are some pre-trained networks developed using ImageNet dataset such that AlexNet [6], VGG-16 [7], GoogLeNet [8], SqueezeNet [9], ResNet [10]. (Table 2 shows the properties of some pre-trained networks). These pre-trained networks have different performances and properties; they are used in transfer learning for image processing. Successful results have been obtained in the studies performed on medical images with these networks trained from natural image dataset [63], [64].

Table 2 Properties of some pre-trained networks

	Year	**Deep**	**Layers**	**MB**	**Architecture**
AlexNet	**2012**	**8**	**25**	**227**	**Traditional sequential network**
VGG-16	**2014**	**16**	**41**	**515**	**Traditional sequential network**
GoogLeNet	**2014**	**22**	**144**	**27**	**Inception module**
ResNet50	**2015**	**50**	**177**	**44**	**Network in network**
Squeezenet	**2016**	**18**	**68**	**4.6**	**Squeeze layer-Fire module**

2.3.10 Fine Tuning

Fine tuning means taking weights from a previously trained neural network and using these weights as a start for a new model to be trained in the same field (usually images). Fine tuning is done to speed up training and overcome small data set size. There are various strategies, such as training using pre-trained weights (usually whole layers). If there is a medium-sized data set to be trained in the hand and is similar to the original data set, the recommended scheme is to use a pre-trained CNN, in which case a successful result is obtained with fine tuning. If the new dataset is large and very different from the original dataset, it would be appropriate to train the network from scratch. In practice, however, it is still very useful to start with weights from a pre-trained network [2].

2.3.11 Hyperparameters

Hyperparameters are the parameters that affect the accuracy and performance of deep neural networks. Hyperparameters set up before training, determine network structure and how the network is trained. Some of these parameters are learning rate, learning rate decay, dropout rate, momentum, early stopping, number of filters, filter size, batch size, and fully connected layers. Some of these are mentioned above. Momentum is a method that helps to speed up the training process with the sgdm approach. Number of epochs is the number of times all training data is shown to the network during training. Batch size is the division of input data into mini lots [5].

2.4 Computer Aided Diagnosis

Medical imaging is a fundamental method in the diagnosis, early diagnosis and differential diagnosis of diseases. Computer Aided Diagnosis (CAD) methods are used to assist physicians in the interpretation of medical images (direct X-rays, MRI images, digital pathology images). In CAD method, the physician first makes a routine evaluation on the image, re-evaluates own interpretation with the help of the CAD system and makes the final decision. Thus the physician receives a second objective opinion aid. Some research suggests that the inclusion of the CAD system in the diagnostic process provides quantitative support for clinical decisions by reducing inter-observer variations [65], [66].

Today, there is a rich data pool with increased resolution as a result of enhancements in computer technology. In addition, image processing has been improving along with advancements in Central Processing Unit (CPU) and Graphics Processing Unit

(GPU) technology. If dataset is big enough, image classification can be successfully performed. Computerized methods, especially image analysis methods, facilitate identification of pathological findings, increase accuracy, and support physician's workflow [48].

Recently, CAD methods has been used in the diagnosis some diseases such as , lung cancer, breast cancer, prostate cancer, colon cancer, bone metastases, coronary artery disease, congenital heart defect, Alzheimer's disease, and diabetic retinopathy [67].

2.5 Related works

Developments in computer technology in recent years have enabled the acquisition and processing of high-resolution images. This development has revolutionized the field of medical image processing and a number of successful studies have been conducted in the field of image recognition. Using CNN method, diagnosis of brain tumors from MRI images, diagnosis of interstitial lung disease using Computed Tomography images and diagnosis of gastric carcinoma by histopathological images are just a few of these successful studies [68], [69], [70].

Plain radiographs have been used in some studies using neural network. Cicero et al studied on chest x-ray and, they used CNN method, and were successful in recognizing various lung pathologies such as normal, pleural effusion, pulmonary edema, consolidation, cardiomegaly, and pneumothorax [71]. In a study performed by Duryea et al., successful results were obtained by using a neural network to determine the localization of joints commonly affected in OA and RA in plain hand radiography [72]. Banerjee et al. were able to detect osteophytes by a cellular neural network method using various filters on hand radiographs, accuracy was over 90% [40]. Murakami et al. for the diagnosis of RA, they used the CNN method to detect bone erosion in the hand joints, their training data set consisted of a total of 129 cases, 90 of which were RA and 39 without RA, testing the performance of their networks with 30 patients with RA, correct positive rate 80.5%, false positive rate was 0.84% [73]. Betancourt-Hernandez et al studied on to diagnose RA using plain hand radiographs, they applied transfer learning using by LeNet, Network in Network and SqueezeNet, and they got successful results [74].

Lee et al, and Spampinato et al., performed two different studies on children to determine of bone age from hand radiographs, they used CNN method, and obtained

successful results. Normally, for each patient, the physician spends about 30 minutes for the bone age detection, they concluded that, using the deep learning approach, the result takes 1 second and its accuracy is better than many conventional methods [61], [75].

There are some studies that have been applied transfer learning on medical images using networks trained with natural images. Shin HC et al., studied on chest and abdomen computed tomography images, they tried to evaluate thoraco-abdominal lymph node and interstitial lung disease classification in their studies. They implemented transfer learning using AlexNet, GoogLeNet, and their modified versions, and achieved successful results [76]. Mednikov Y et al., studied on mammography images for breast mass detection and applied transfer learning using inception-v3 network. They used the INBreast database, produced artificial mammograms, and used different augmentation techniques, and obtained maximal area of the receiver operating characteristics curve of 0.91 [77].

Another group, Huynh BQ., et al worked with mammography images for tumor classification and applied transfer learning using AlexNet. They used dataset containing 219 breast lesions to distinguish benign and malignant breast lesions. They achieved good performance with transfer learning (area under the ROC curve (AUC) = 0.81) [63]. Cicero M et al studied on chest radiographs, applied transfer learning using GoogLeNet and, they found an overall sensitivity and specificity of 91% [71].

Kim DH, MacKinnon T., tried to diagnose hand fractures on plain hand radiographs, applied transfer learning using inception-v3 network, found sensitivity and specificity values of 0.9 and 0.88, respectively. They concluded that transfer learning using CNN pre-trained with non-medical images, could be applied to analysis of plain radiographs [78].

Xue Y et al. studied on hip osteoarthritis. Their dataset consisted of 420 pelvic radiographs. They applied transfer learning with VGG16 network, they achieved sensitivity of 95.0%, and specificity of 90.7% [79]. Tiulpin A et al. studied patients with knee osteoarthritis and performed fine tuning with ResNet34 network. They studied on 5,960 knees radiography and obtained 66.71% of average multiclass accuracy [80].

CHAPTER 3

MATERIALS and METHODS

This section contains information about data acquisition, classification, preprocessing operations, number of data used for testing and training, pre-trained networks used in this study, and hyperparameters applied during data training.

3.1 Dataset

This study was carried out at the Medical Faculty of Kırıkkale University, Rheumatology Department. Plain hand radiographs of patients who were examined in rheumatology outpatient clinic between 1 January 2012 and 1 May 2019 were used as dataset. Ethical approval of this study was obtained from the Kırıkkale University Non-Interventional Research Ethics Committee on Sep 12th 2018, no. 2018.09.01-02. There were some artifacts on the images, such as rings, keys, medical supplies; these images were not taken into the study (Figure 11 shows noisy images samples).

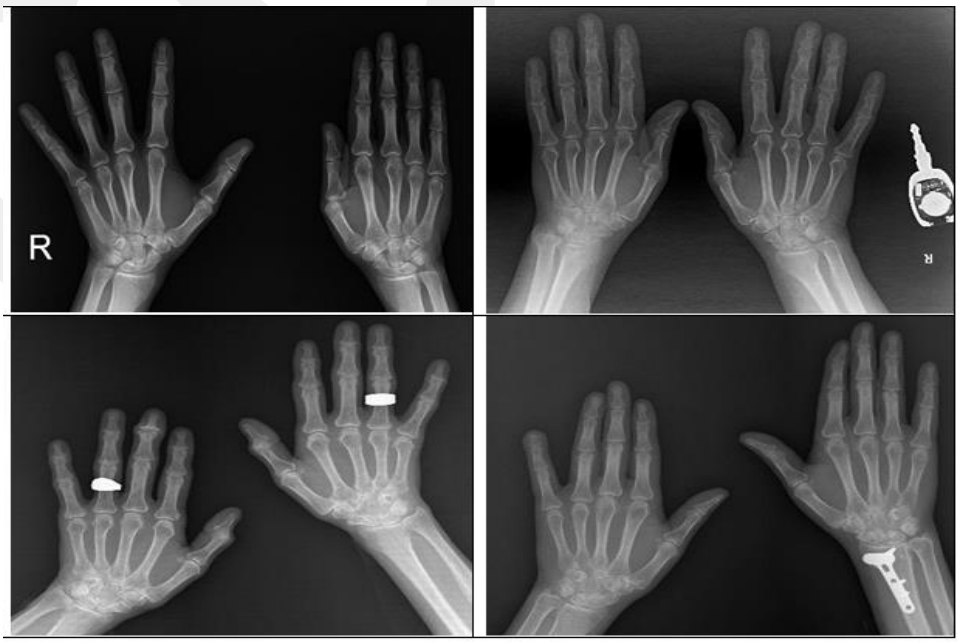


Figure 11 Samples of noisy images

Conventional hand radiographs (CR) taken in Posteroanterior position have been classified as normal, rheumatoid arthritis and hand osteoarthritis radiographs (first classification). Radiographs classified as RA and OA by the first observer (writer) were then classified by another two experienced rheumatologists without being aware of each other, and radiography that was not accepted as OA or RA by one of these rheumatologists was not included in the study (second classification). (Prof. Dr. Abdurrahman Tufan, Gazi University, Faculty of Medicine, Department of Rheumatology and Assoc. Prof. Dr. Levent Kılıç, Hacettepe University, Faculty of Medicine, Department of Rheumatology). (The number of hand radiographs obtained after the first and second classifications are shown in Table 3.)

Table 3 Number of hand radiographs obtained after the first and second classifications

	First classification	**Second classification**
Rheumatoid arthritis	**190**	**152**
Osteoarthritis	**178**	**146**
Normal	**158**	**150**

3.2 Data Pre-processing and Splitting

Patients archive images were in jpg format, with widths ranging from 500 and 600, heights between 600 and 900 pixels and resolution 96 dpi. There were some artifacts on the images, such as rings, keys, medical supplies. For this reason, the right and left hands were cropped separately so that irrelevant features would not affect the training, and to increase the dataset. White color padding was added to the left and right sides of each image to prevent data loss and distortion and to obtain a square frame, thus the resolution did not change. And then, each image was resized to 224 x 224 x 3 dimensions. Finally, left hand flipped. As a result, the dataset consisted of 300 normal hand radiographic images, 304 radiographic images of RA and 292 radiographic images of hand OA. (Figure 12 shows some cropped and resized images samples)

Figure 12 Cropped and resized images samples

Apart from RA and OA, diseases such as psoriatic arthritis, gout, calcium pyrophosphate arthritis, scleroderma, and fractures also cause changes in hand radiographs. The inclusion of an image with an incorrect tag is not conducive to training an automated system. Kim DH, and MacKinnon T, who worked on hand fractures, proposed to include a third diagnostic category of "inconclusive" to automated systems to cope with uncertainty [78]. In this study, a 4th class was added called "others". Distal radius fractures radiographs, foot and knee radiographs, pelvis radiographs, chest radiographs were added to this class and the same preprocessing procedures were performed. (Figure 13 shows some sample images in "others" class)

Figure 13 Others class samples

As a result, the dataset consisted of 300 normal hand radiographic images, 304 hand radiographic images of RA, and 292 hand radiographic images of OA, while others class had 240 images. These radiographs were randomly split into two as training and test data, 80% were used as training data and 20% were used as test data.(Table 3 shows the number of training and validation images). Table 4 shows this splitting.

Table 4 Number of training and validation images

	Training	Validation	Total
Normal	**240**	**60**	**300**
Rheumatoid arthritis	**243**	**61**	**304**
Osteoarthritis	**234**	**58**	**292**
other	**192**	**48**	**240**

3.3 Data Augmentation

Data augmentation was done by rotating images randomly, rotating the images horizontally and vertically by 3 pixels and rotating the images up to 30 degrees.

3.4 Transfer Learning

In this study, pre-trained networks VGG16, ResNet50 and GoogLeNet were used for transfer learning. First, the original classifier was removed and a binary classifier was added. Same hyperparameters were used in each model. (Table 5 shows the hyperparameters used in this study).

Table 5 Hyperparameters used in this study

Optimizer	**sgdm**
MiniBatchSize	**16**
InitialLearnRate	**1e-4**
LearnRateDropFactor	**0.2**
LearnRateDropPeriod	**8**
L2Regularization	**0.004**
ValidationFrequency	**16**

CHAPTER 4

RESULTS

This work was performed on a LENOVO Intel® Core ™ i7-9750H / Y540 / 16G / 512 GeForce RTX2060 computer and in MATLAB environment. The efficiency of the created models was evaluated by using performance metrics such as accuracy, sensitivity, specificity and precision. These performance metrics were calculated using the confusion matrices. (Figure 14-17 show training charts of VGG16, Table 6-9 show confusion matrices of VGG16, Figure 18-21 show training charts of ResNet50, Table 10-13 show confusion matrices of ResNet50, Figure 22-25 show training charts of GoogLeNet, Table 14-17 show confusion matrices of GoogLeNet) (After training using VGG16, ResNet50, and GoogLeNet obtained accuracy, precision, sensitivity, specificity results calculated using equations(2), (3), (4) and (5) are shown in Table 18)

$$Accuracy = \frac{(TP+TN)}{(TP+TN+FP+FN)} \quad (2)$$

TP: True Positive, TN: True Negative, FP: False Positive, FN: False Negative

$$Sensitivity = \frac{TP}{(TP+FN)} \quad (3)$$

$$Specificity = \frac{TN}{(TN+FP)} \quad (4)$$

$$Precision = \frac{TP}{(TP+FP)} \quad (5)$$

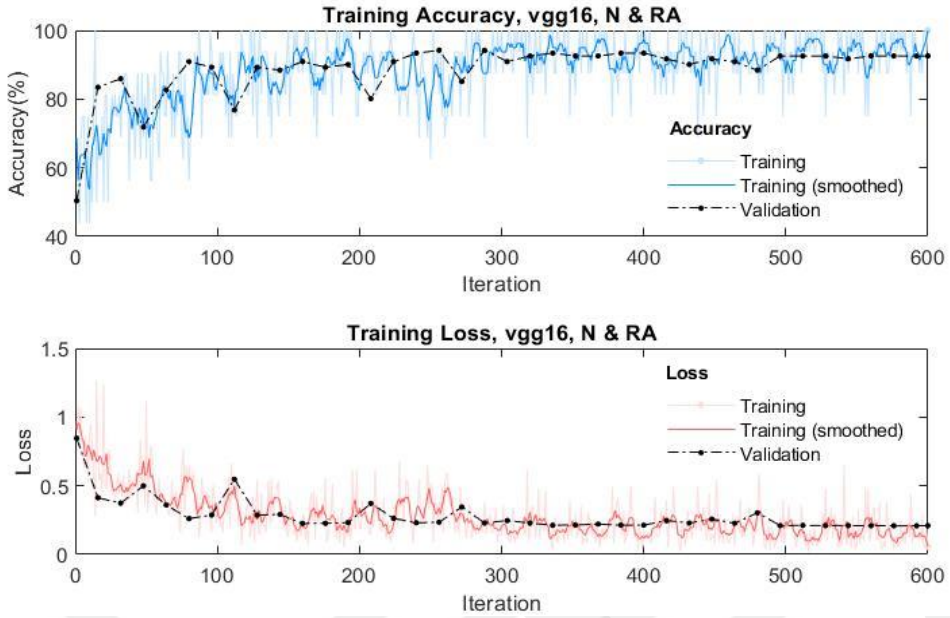


Figure 14 VGG16 training chart, N & RA

Table 6 VGG16 confusion matrix, N & RA

Confusion matrix (VGG16)				
Output class	romato	54	7	88.5%
	normal	2	58	96.7%
		96.4%	89.2%	92.6 %
		romato	normal	
	Target class			

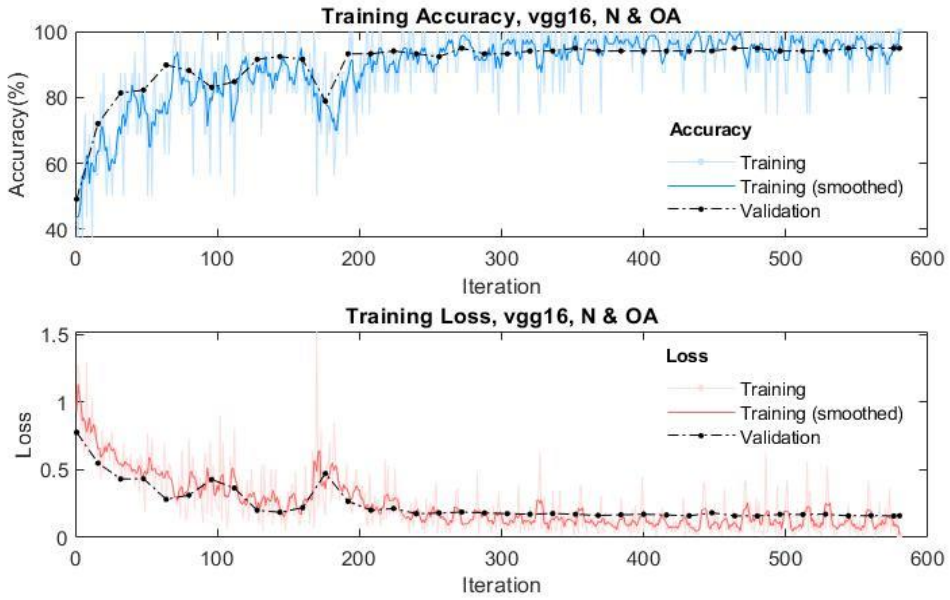


Figure 15 VGG16 training chart, N & OA

Table 7 VGG16 confusion matrix, N & OA

Confusion matrix (VGG16)				
Output class	osteo	54	4	93.1%
	normal	2	58	96.7%
		96.4%	93.5%	94.9 %
		osteo	normal	
	Target class			

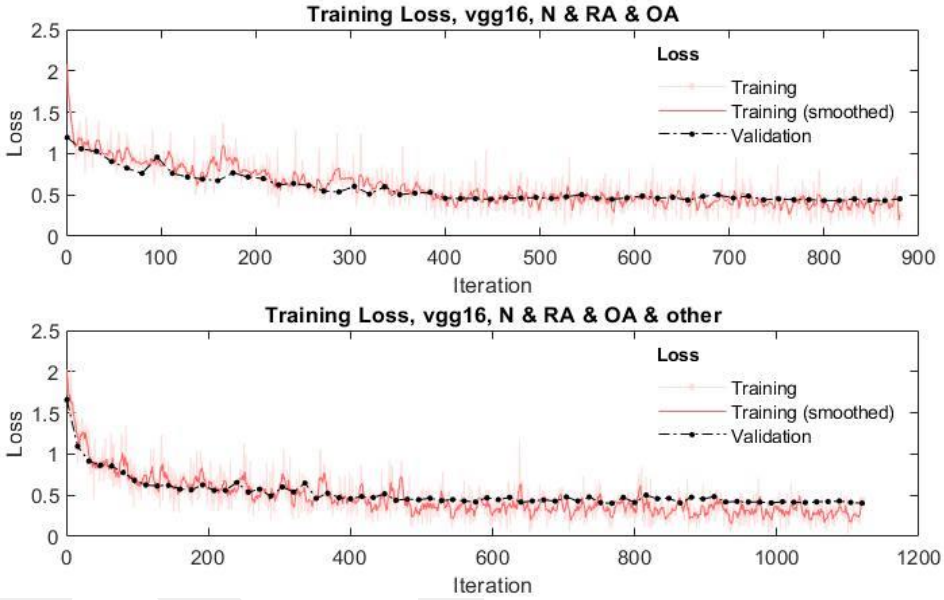


Figure 16 VGG16 training chart, N & RA & OA

Table 8 VGG16 confusion matrix training chart, N & RA & OA

	Confusion matrix (VGG16)				
Output class	osteo	45	11	2	77.6%
	romato	4	52	5	85.2%
	normal	1	5	54	90.0%
		90.0%	76.5%	88.5%	84.4%
		osteo	romato	normal	
	Target class				

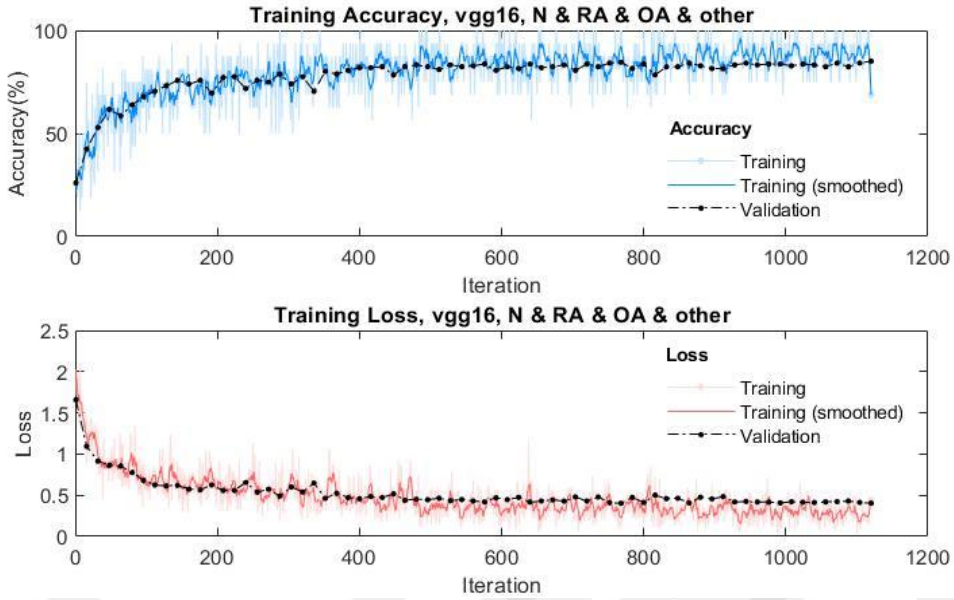


Figure 17 VGG16 training chart, N & RA & OA & other

Table 9 VGG16 confusion matrix, N & RA & OA & other

	Confusion matrix (VGG16)					
Output class	osteo	42	11	5	0	72.4%
	romato	0	46	6	0	75.4%
	normal	2	1	57	0	95.0%
	other	0	0	0	48	100%
		79.2%	79.3%	83.8%	100%	85.0%
		osteo	romato	Normal	other	
	Target class					

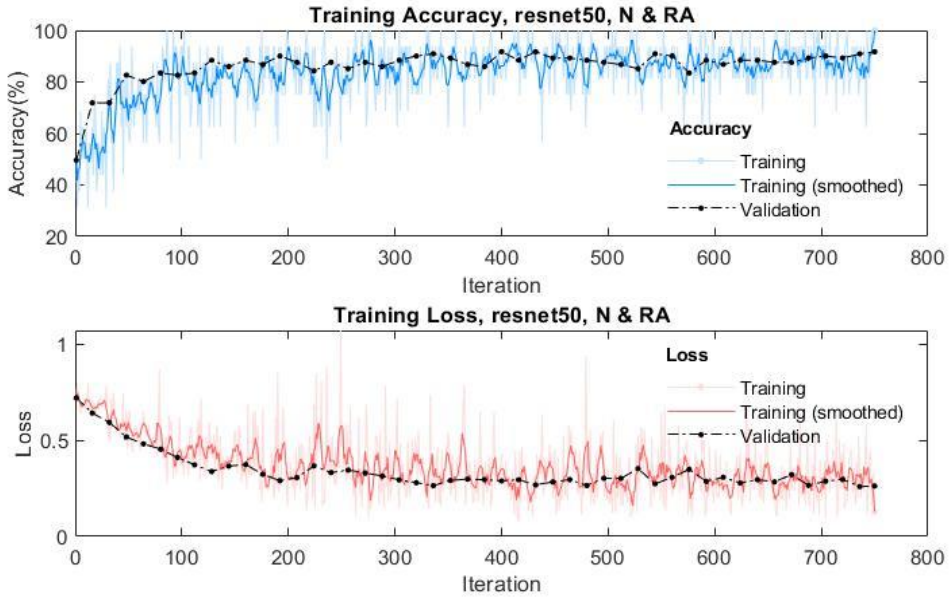


Figure 18 ResNet50 training chart, N & RA

Table 10 ResNet50 confusion matrix, N & RA

Confusion matrix (ResNet50)				
Output class	romato	56	5	91.8%
	normal	5	55	91.7%
		91.8%	91.7%	91.7 %
		romato	normal	
	Target class			

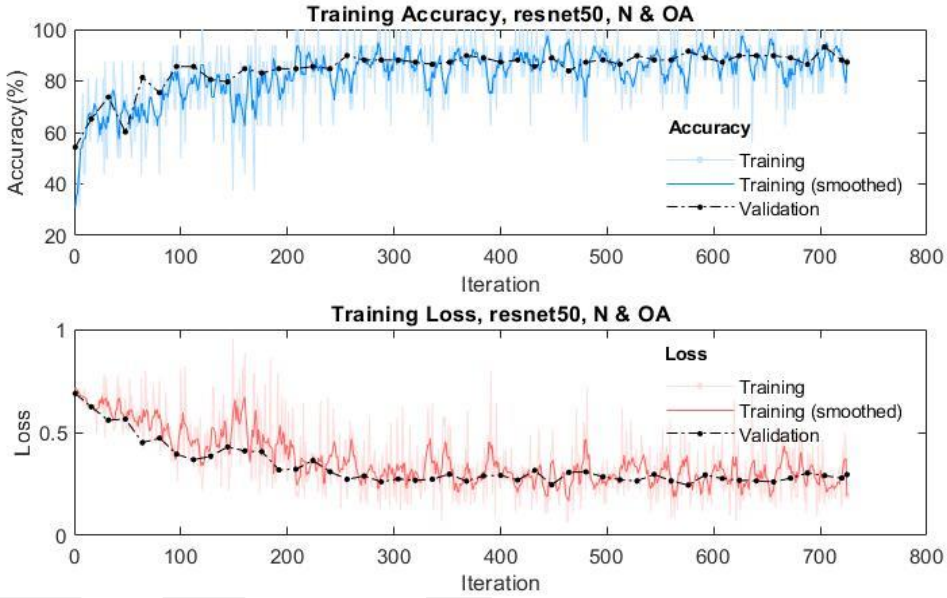


Figure 19 ResNet50 training chart, N & OA

Table 11 ResNet50 confusion matrix, N & OA

Confusion matrix (ResNet50)				
Output class	osteo	54	4	91.3%
	normal	4	56	93.3%
		93.1%	93.3%	93.2 %
		osteo	normal	
	Target class			

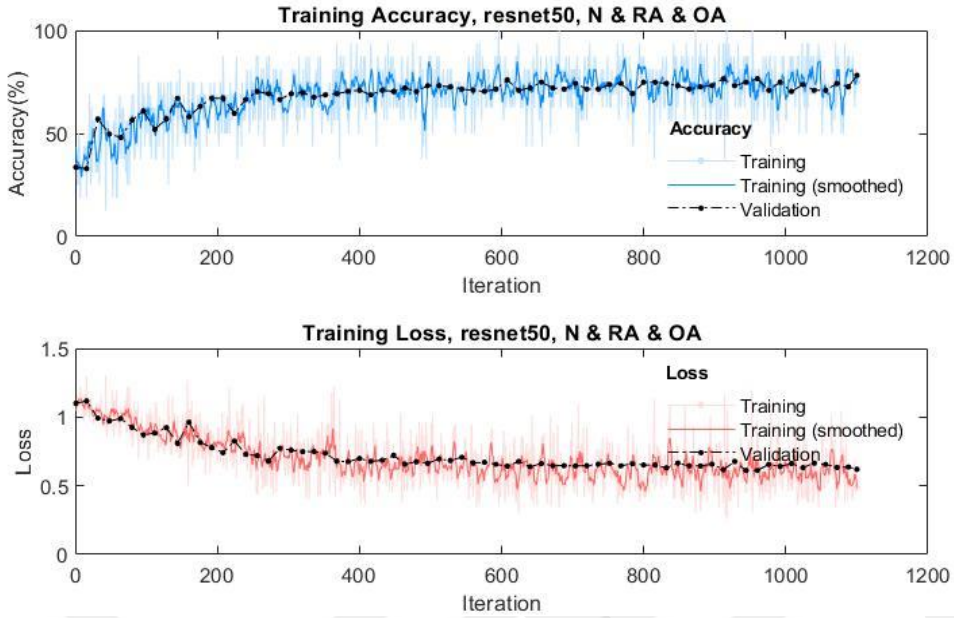


Figure 20 ResNet50 training chart, N & RA & OA

Table 12 ResNet50 confusion matrix, N & RA & OA

	Confusion matrix (ResNet50)				
Output class	Osteo	30	24	4	51.7%
	romato	7	50	4	82.0%
	normal	0	6	54	90.0%
		81.1%	62.5%	87.1%	74.9%
		osteo	romato	normal	
	Target class				

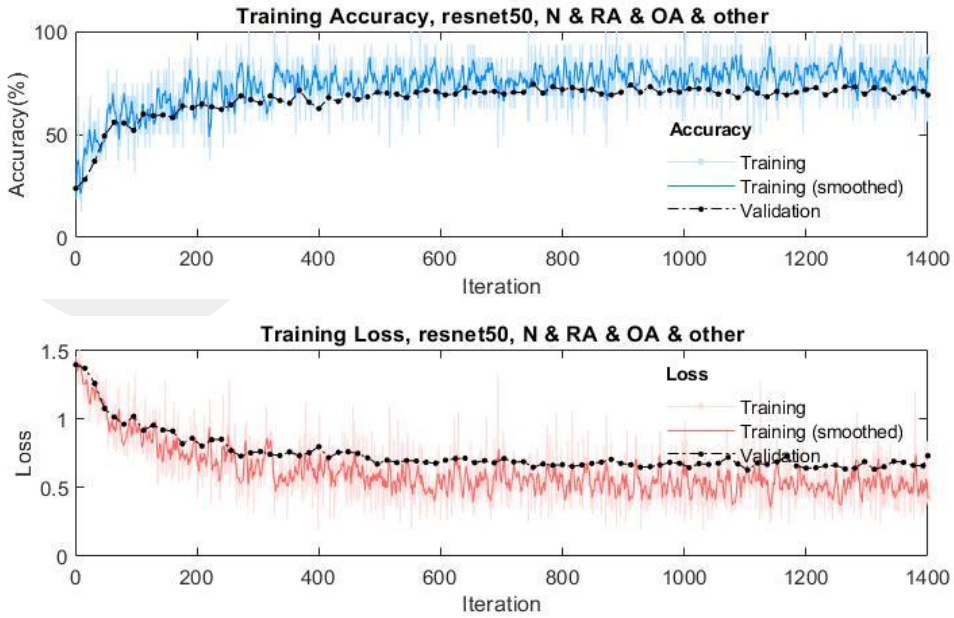


Figure 21 ResNet50 training chart, N & RA & OA & other

Table 13 ResNet50 confusion matrix, N & RA & OA & other

	Confusion matrix (ResNet50)					
Output class	osteo	35	17	6	0	60.3%
	romato	12	43	6	0	70.5%
	normal	4	3	53	0	88.3%
	other	0	0	2	46	95.8%
		68.6%	68.3%	79.1%	100%	78.0%
		osteo	Romato	normal	other	
	Target class					

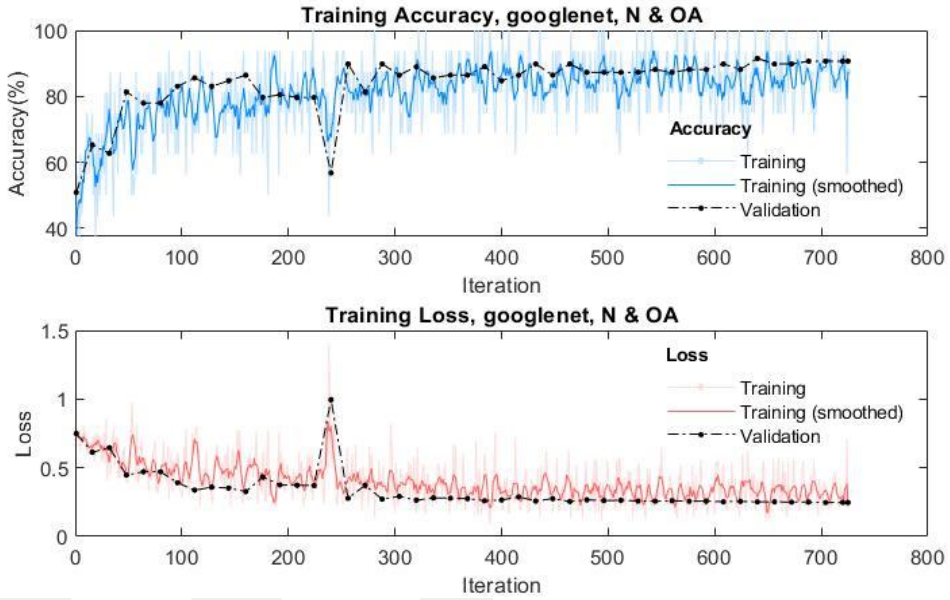


Figure 22 GoogLeNet training chart, N & OA

Table 14 GoogLeNet confusion matrix, N & OA

Confusion matrix (GoogLeNet)				
Output class	osteo	50	8	86.2%
	normal	3	57	95.0%
		94.3%	87.7%	90.7 %
		osteo	normal	
	Target class			

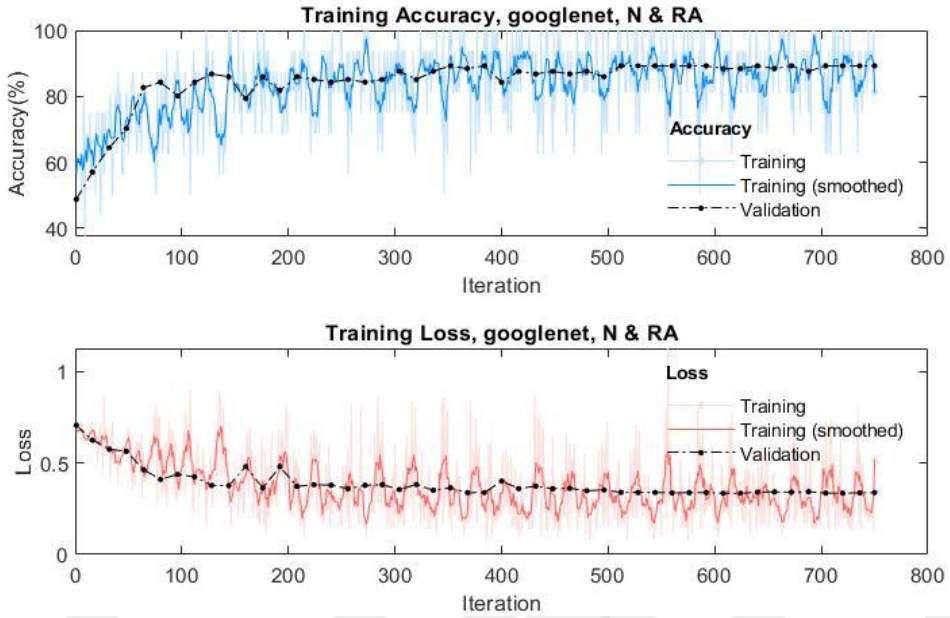


Figure 23 GoogLeNet training chart, N & RA

Table 15 GoogLeNet confusion matrix, N & RA

Confusion matrix (GoogLeNet)				
Output class	romato	56	5	91.8%
	normal	8	52	86.7%
		87.5%	91.2%	89.3 %
		romato	normal	
	Target class			

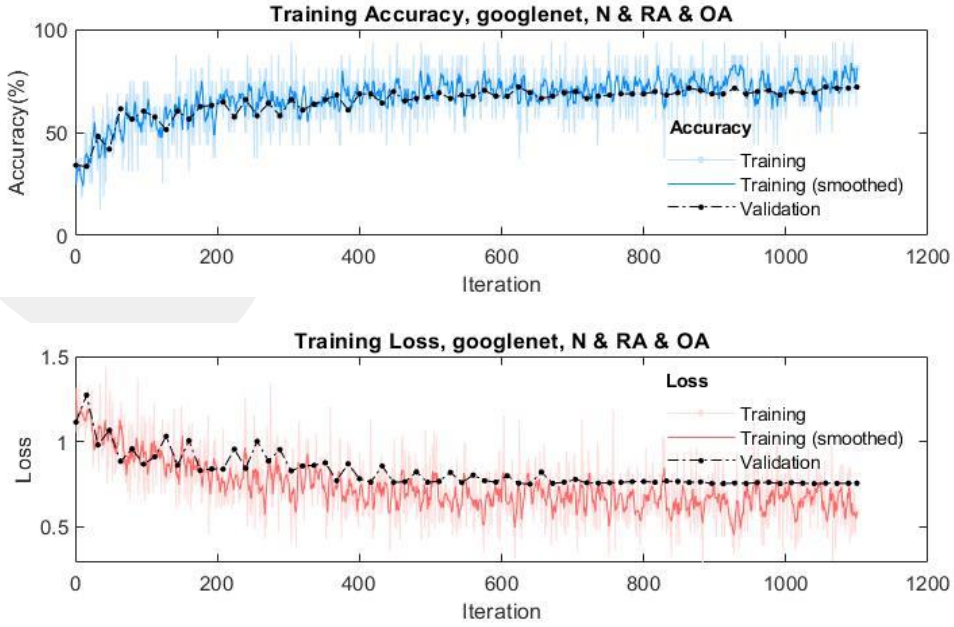


Figure 24 GoogLeNet training chart, N & OA & RA

Table 16 GoogLeNet confusion matrix, N & OA & RA

	Confusion matrix (GoogLeNet)				
Output class	osteo	34	14	10	58.6%
	romato	10	46	5	75.4%
	normal	5	6	49	81.7%
		69.4%	69.7%	76.6%	72.1%
		osteo	romato	normal	
	Target class				

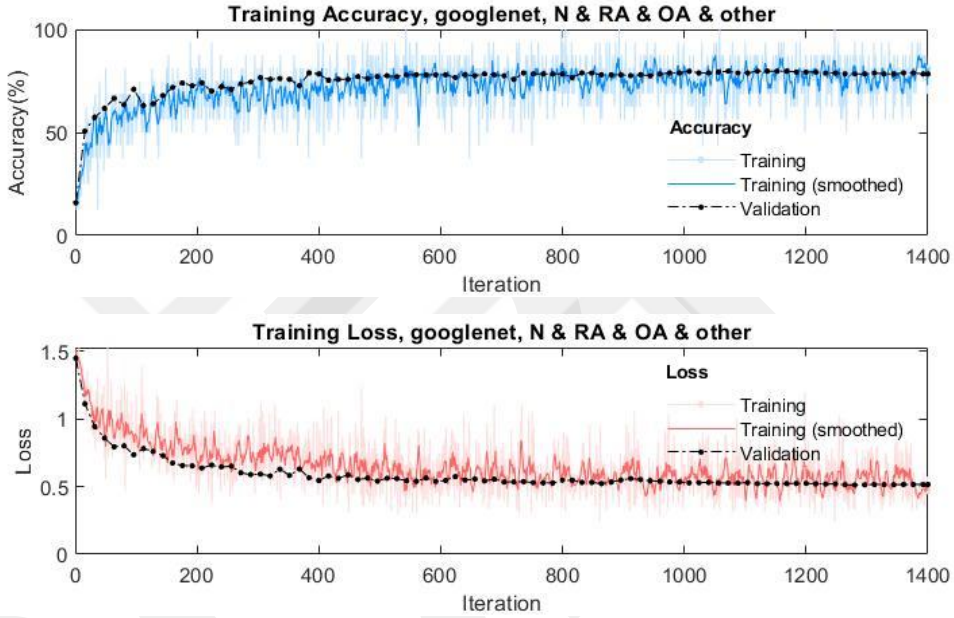


Figure 25 GoogLeNet training chart, N & OA & RA & other

Table 17 GoogLeNet confusion matrix, N & OA & RA & other

	Confusion matrix (GoogLeNet)					
Output class	osteo	27	23	8	0	46.6%
	romato	4	50	7	0	82.0%
	normal	2	5	53	0	88.3%
	other	0	0	0	48	100%
		81.8%	64.1%	77.9%	100%	78.4%
		osteo	romato	normal	other	
	Target class					

Table 18 Training results of VGG16, ResNet50 and GoogLeNet

			VGG16	ResNet50	GoogLeNet
N & RA	**Accuracy**		**92.6**	**91.7**	**89.3**
	Precision		**88.5**	**91.8**	**91.8**
	Sensitivity		**96.4**	**91.8**	**87.5**
	Specificity		**89.2**	**91.7**	**91.2**
N & OA	**Accuracy**		**94.9**	**93.2**	**90.7**
	Precision		**93.1**	**93.1**	**86.2**
	Sensitivity		**96.4**	**93.1**	**94.3**
	Specificity		**93.5**	**93.3**	**87.7**
N & RA & OA	**Accuracy**		**84.4**	**74.9**	**72.1**
	Precision	**RA**	**85.2**	**82.0**	**75.4**
		OA	**77.6**	**51.7**	**58.6**
	Sensitivity	**RA**	**76.5**	**62.5**	**69.4**
		OA	**90.0**	**81.1**	**69.7**
N & RA & OA & other	**Accuracy**		**85.0**	**78.0**	**78.4**
	Precision	**RA**	**75.4**	**70.5**	**82.0**
		OA	**72.4**	**60.3**	**46.6**
	Sensitivity	**RA**	**79.3**	**68.3**	**64.1**
		OA	**79.2**	**68.6**	**81.8**

CHAPTER 5

DISCUSSIONS & CONCLUSION

5.1 Discussions

Several studies have been performed by using CNN on knee and hip OA using plain radiographs [4], [79], [80]. According to my knowledge, our study is the first study on hand OA using CNN and transfer learning. Murakami et al., for the diagnosis of RA, they used the CNN method to detect bone erosion in the hand joints, their training data set consisted of a total of 129 cases, 90 of which were RA and 39 without RA, testing the performance of their networks with 30 patients with RA, correct positive rate 80.5%, false positive rate was 0.84% [73]. Betancourt-Hernandez et al studied on to diagnose RA using plain hand radiographs, they applied transfer learning using by LeNet, Network in Network and SqueezeNet, and they got successful results [74].

To the best our knowledge, this is the first study to distinguish between normal, hand OA and RA using plain hand radiographs. In this study, pre-trained GoogLeNet, ResNet50 and VGG-16 networks were used, transfer learning was applied. Successful results were obtained with all three pre-trained networks in differentiation of normal and RA hand x-ray images and in differentiation of normal and OA hand x-ray images. Similarly, promising results were obtained with all three networks in normal, OA and RA hand x-ray images and differentiation of normal, OA, RA and "other" class images.

CNN training from scratch requires a lot of data, transfer learning can be applied when there is not enough data. Transfer learning is the use of a pre-trained model on a new problem. It is particularly difficult to find sufficient data in the medical field, in this case, CNN models previously trained from a natural image data set or from a different medical field are used for a new medical task [5]. In recent years, there have been a few studies showing that CNN models trained with natural images have achieved successful results in the medical field. This study supports previous studies showing that pre-trained modals trained with non-medical images can be successfully used in the classification of medical images.

Radiological findings of RA and OA occur relatively late, but as mentioned above, in some patients with RA or OA, the patient's history and laboratory tests are not sufficiently useful for the diagnosis, in which case the imaging method becomes more important. The potential benefit of this method we are trying to improve is that it can be used in these patients. If RA or OA is diagnosed using CR, further imaging methods such as MRI or ultrasonography are not necessary. MRI is an expensive and time consuming imaging method and ultrasonography is an operator dependent imaging method.

Hand OA and RA are two different diseases that cause pain, swelling, tenderness, loss of function in hand joints. Treatments of both diseases are also different. A patient with pain, swelling, or stiffness in the hand joints may be examined by a general practitioner, internal medicine specialist, orthopedic specialist, physical therapy specialist, or rheumatologist. The physician's experience in assessing RA and OA findings on the CR is important. There may not be experienced physicians in any center where CR is available. Physicians who do not have sufficient experience to evaluate of CR can differentiate RA and OA using this CAD method which we are trying to develop. In this case, RA patients are referred to rheumatologists for early treatment. With early RA treatment, permanent deformities seen in RA can be prevented. And with early diagnosis of OA, patient's quality of life and improvement in the course of disease are improved by simple procedures such as patient education and lifestyle modification. Thus, doctors working in primary health care institutions may take an active role in RA and OA treatment.

In addition, experienced physicians, such as rheumatologists and radiologists, can use this CAD method that we are trying to develop when making final decisions about hand radiographs. Thus, unwanted results that may arise due to work intensity, fatigue, carelessness, insufficient time may be prevented, physician receives objective second opinion, and the concern of inter-observer and intra-observer reliability can be reduced.

During the testing of a trained network, assigns each image to one of the classes in which it is trained. In other words, the network trained classification task for normal hand radiography, hand radiography with OA, and hand radiography with RA, determines each image in one of these three classes during the test. Apart from RA and OA, diseases such as psoriatic arthritis, gout, calcium pyrophosphate arthritis, scleroderma, and fractures also cause changes in hand radiographs. In this study, a

4th class was added called "others". Distal radius fractures, phalangeal fractures, scleroderma hand radiographs, foot images, pelvis radiographs, chest radiographs were added to this class and the same preprocessing procedures were performed.

The addition of the "others" class as the 4th class has made the model robust and made it suitable for outpatient clinics. In other words, for radiographic evaluation, when a radiography is presented to this model other than normal hand, RA and OA patient hand radiography, the result will be "others". Likewise, if a radiography of other sites such as foot radiography, lung radiography is presented to the model, the result will be "others". The addition of the "others" class has made the model also expandable. If enough radiography is collected in a disease class such as gout or calcium pyrophosphate arthritis, a new class is created as 5th class, and so on.

In the CNN method, overfitting is an important problem to be tackled. In this study, data augmentation, droupout, fine tuning, learning rate decay were applied to prevent overfitting, and no signs of overfitting were observed in the training chart, the training accuracy graph and the validation accuracy graph were close to each other. There was no training accuracy and validation loss mismatch.

5.2 Conclusion

This study was performed using CRs of patients examined in a single rheumatology center, although data augmentation and transfer learning were applied, the number of patients for CNN training should be much higher, however, the results of this preliminary study are promising. New studies can be conducted in cooperation with several centers to increase the number of patients. Thus, this model can be improved and sufficient radiographs of other hand-related diseases can be collected and the model can be developed and expanded.

Our patients had early and late radiological changes due to RA on their radiographs, but we did not calculate the Modified Sharp score. Likewise, we did not calculate disease activity scores in patients with OA, such as Kellgren-Lawrence. If this could be done, it would be seen how successful the model was in the diagnosis of early RA and early OA. Further work is needed in this regard.

Radiographic changes in RA and OA occur relatively late compared with MRI and ultrasonography findings. Another research area is the studies on CRs of patients diagnosed with MRI and ultrasonography. Because the changes in the pixel level that cannot be seen by the human eye can be detected by deep learning methods.

REFERENCES

[1] Y. LeCun, L. Bottou, Y. Bengio, P. H.-P. of the IEEE, and undefined 1998, "Gradient-based learning applied to document recognition," *dengfanxin.cn*.

[2] H. Greenspan, … B. V. G.-I. T. on, and undefined 2016, "Guest editorial deep learning in medical imaging: Overview and future promise of an exciting new technique," *ieeexplore.ieee.org*.

[3] E. C. Too, L. Yujian, S. Njuki, and L. Yingchun, "A comparative study of fine-tuning deep learning models for plant disease identification," *Comput. Electron. Agric.*, Mar. 2018.

[4] G. Litjens *et al.*, "A survey on deep learning in medical image analysis," *Medical Image Analysis*, vol. 42. Elsevier B.V., pp. 60–88, 01-Dec-2017.

[5] S. M. Anwar, M. Majid, A. Qayyum, M. Awais, M. Alnowami, and M. K. Khan, "Medical Image Analysis using Convolutional Neural Networks: A Review," *J. Med. Syst.*, vol. 42, no. 11, p. 226, Nov. 2018.

[6] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "ImageNet Classification with Deep Convolutional Neural Networks."

[7] K. Simonyan and A. Zisserman, "Very Deep Convolutional Networks for Large-Scale Image Recognition," Sep. 2014.

[8] C. Szegedy *et al.*, "Going Deeper with Convolutions," Sep. 2014.

[9] F. N. Iandola, S. Han, M. W. Moskewicz, K. Ashraf, W. J. Dally, and K. Keutzer, "SqueezeNet: AlexNet-level accuracy with 50x fewer parameters and," Feb. 2016.

[10] K. He, X. Zhang, S. Ren, and J. Sun, "Deep Residual Learning for Image Recognition." pp. 770–778, 2016.

[11] C. Szegedy, S. Ioffe, V. Vanhoucke, A. A.-T.-F. A. Conference, and undefined 2017, "Inception-v4, inception-resnet and the impact of residual connections on learning," *aaai.org*.

[12] S. E. Gabriel, "The epidemiology of rheumatoid arthritis," *Rheum. Dis. Clin. North Am.*, vol. 27, no. 2, pp. 269–281, 2001.

[13] V. Picerno, F. Ferro, A. Adinolfi, E. Valentini, C. Tani, and A. Alunno, "One year in review: the pathogenesis of rheumatoid arthritis.," *Clin. Exp.*

Rheumatol., vol. 33, no. 4, pp. 551–8.

[14] D. L. Scott, F. Wolfe, and T. W. J. Huizinga, "Rheumatoid arthritis.," *Lancet (London, England)*, vol. 376, no. 9746, pp. 1094–108, Sep. 2010.

[15] A. Fleming, J. M. Crown, and M. Corbett, "Early rheumatoid disease. I. Onset," *Ann. Rheum. Dis.*, vol. 35, no. 4, pp. 357–360, 1976.

[16] R. K. Jacoby, M. I. Jayson, and J. A. Cosh, "Onset, early stages, and prognosis of rheumatoid arthritis: a clinical study of 100 patients with 11-year follow-up.," *Br. Med. J.*, vol. 2, no. 5858, pp. 96–100, Apr. 1973.

[17] S. Lineker, E. Badley, C. Charles, L. Hart, and D. Streiner, "Defining morning stiffness in rheumatoid arthritis.," *J. Rheumatol.*, vol. 26, no. 5, pp. 1052–7, May 1999.

[18] S. E. Gabriel, J. L. Wagner, A. R. Zinsmeister, C. G. Scott, and H. S. Luthra, "Is rheumatoid arthritis care more costly when provided by rheumatologists compared with generalists?," *Arthritis Rheum.*, vol. 44, no. 7, pp. 1504–1514, 2001.

[19] S. Vyas, A. S. Bhalla, P. Ranjan, S. Kumar, U. Kumar, and A. K. Gupta, "Rheumatoid Arthritis Revisited – Advanced Imaging Review," *Polish J. Radiol.*, vol. 81, pp. 629–635, Dec. 2016.

[20] M. Kourilovitch, C. Galarza-Maldonado, and E. Ortiz-Prado, "Diagnosis and classification of rheumatoid arthritis.," *J. Autoimmun.*, vol. 48–49, pp. 26–30, Feb. 2014.

[21] D. M. Lee and M. E. Weinblatt, "Rheumatoid arthritis.," *Lancet (London, England)*, vol. 358, no. 9285, pp. 903–11, Sep. 2001.

[22] W. R. Renner and A. S. Weinstein, "Early changes of rheumatoid arthritis in the hand and wrist.," *Radiol. Clin. North Am.*, vol. 26, no. 6, pp. 1185–93, Nov. 1988.

[23] J. T. Sharp *et al.*, "How many joints in the hands and wrists should be included in a score of radiologic abnormalities used to assess rheumatoid arthritis?," *Arthritis Rheum.*, vol. 28, no. 12, pp. 1326–35, Dec. 1985.

[24] J. F. Baker, P. G. Conaghan, and F. Gandjbakhch, "Update on magnetic resonance imaging and ultrasound in rheumatoid arthritis.," *Clin. Exp. Rheumatol.*, vol. 36 Suppl 114, no. 5, pp. 16–23.

[25] E. Niemantsverdriet and A. H. M. van der Helm-van Mil, "Imaging detected tenosynovitis of metacarpophalangeal and wrist joints: an increasingly

recognised characteristic of rheumatoid arthritis.," *Clin. Exp. Rheumatol.*, vol. 36 Suppl 1, no. 5, pp. 131–138.

[26] B. J. Harrison, D. P. Symmons, E. M. Barrett, and A. J. Silman, "The performance of the 1987 ARA classification criteria for rheumatoid arthritis in a population based cohort of patients with early inflammatory polyarthritis. American Rheumatism Association.," *J. Rheumatol.*, vol. 25, no. 12, pp. 2324–30, Dec. 1998.

[27] J. S. Smolen *et al.*, "EULAR recommendations for the management of rheumatoid arthritis with synthetic and biological disease-modifying antirheumatic drugs: 2013 update," *Ann. Rheum. Dis.*, vol. 73, no. 3, pp. 492–509, Mar. 2014.

[28] G. R. Burmester and J. E. Pope, "Novel treatment strategies in rheumatoid arthritis," *The Lancet*, vol. 389, no. 10086. Lancet Publishing Group, pp. 2338–2348, 10-Jun-2017.

[29] C. A. *et al.*, "One year in review 2016: Novelties in the treatment of rheumatoid arthritis," *Clin. Exp. Rheumatol.*, vol. 34, no. 3, pp. 357–372, 2016.

[30] G. J. Leung, K. D. Rainsford, and W. F. Kean, "Osteoarthritis of the hand I: aetiology and pathogenesis, risk factors, investigation and diagnosis," *J. Pharm. Pharmacol.*, vol. 66, no. 3, pp. 339–346, Mar. 2014.

[31] Y. Zhang and J. M. Jordan, "Epidemiology of osteoarthritis.," *Clin. Geriatr. Med.*, vol. 26, no. 3, pp. 355–69, Aug. 2010.

[32] A. Mobasheri and M. Batt, "An update on the pathophysiology of osteoarthritis," *Annals of Physical and Rehabilitation Medicine*, vol. 59, no. 5–6. Elsevier Masson SAS, pp. 333–339, 01-Dec-2016.

[33] M. Rossignol *et al.*, "Primary osteoarthritis of hip, knee, and hand in relation to occupational exposure.," *Occup. Environ. Med.*, vol. 62, no. 11, pp. 772–7, Nov. 2005.

[34] G. J. Leung, K. D. Rainsford, and W. F. Kean, "Osteoarthritis of the hand I: aetiology and pathogenesis, risk factors, investigation and diagnosis.," *J. Pharm. Pharmacol.*, vol. 66, no. 3, pp. 339–46, Mar. 2014.

[35] B. Xia, Di Chen, J. Zhang, S. Hu, H. Jin, and P. Tong, "Osteoarthritis pathogenesis: a review of molecular mechanisms.," *Calcif. Tissue Int.*, vol. 95, no. 6, pp. 495–505, Dec. 2014.

[36] "The American College of Rheumatology criteria for the classification and reporting of osteoarthritis of the hand," *Arthritis Rheum.*, vol. 33, no. 11, pp. 1601–1610, 1990.

[37] K. D. Allen, C. J. Coffman, Y. M. Golightly, K. M. Stechuchak, and F. J. Keefe, "Daily pain variations among patients with hand, hip, and knee osteoarthritis," *Osteoarthr. Cartil.*, vol. 17, no. 10, pp. 1275–1282, Oct. 2009.

[38] A. Mathiessen, M. A. Cimmino, H. B. Hammer, I. K. Haugen, A. Iagnocco, and P. G. Conaghan, "Imaging of osteoarthritis (OA): What is new?," *Best Pract. Res. Clin. Rheumatol.*, vol. 30, no. 4, pp. 653–669, 2016.

[39] F. W. Roemer, F. Eckstein, D. Hayashi, and A. Guermazi, "The role of imaging in osteoarthritis," *Best Practice and Research: Clinical Rheumatology*, vol. 28, no. 1. Bailliere Tindall Ltd, pp. 31–60, 2014.

[40] S. Banerjee, S. Bhunia, and G. Schaefer, "Osteophyte detection for hand osteoarthritis identification in x-ray images using CNNs," in *Proceedings of the Annual International Conference of the IEEE Engineering in Medicine and Biology Society, EMBS*, 2011, pp. 6196–6199.

[41] M. T. Nieminen, V. Casula, M. T. Nevalainen, and S. Saarakkala, "Osteoarthritis year in review 2018: imaging.," *Osteoarthr. Cartil.*, vol. 27, no. 3, pp. 401–411, Mar. 2019.

[42] R. Ramonda, P. Frallonardo, E. Musacchio, S. Vio, and L. Punzi, "Joint and bone assessment in hand osteoarthritis," *Clinical Rheumatology*, vol. 33, no. 1. pp. 11–19, Jan-2014.

[43] J. Hill and H. Bird, "Patient knowledge and misconceptions of osteoarthritis assessed by a validated self-completed knowledge questionnaire (PKQ-OA).," *Rheumatology (Oxford).*, vol. 46, no. 5, pp. 796–800, May 2007.

[44] D. Pereira, E. Ramos, and J. Branco, "Osteoarthritis.," *Acta Med. Port.*, vol. 28, no. 1, pp. 99–106.

[45] M. Shaban *et al.*, "Automated Staging of Diabetic Retinopathy Using a 2D Convolutional Neural Network," in *2018 IEEE International Symposium on Signal Processing and Information Technology, ISSPIT 2018*, 2019, pp. 354–358.

[46] Y. Lecun, Y. Bengio, and G. Hinton, "Deep learning," *Nature*, vol. 521, no. 7553. Nature Publishing Group, pp. 436–444, 27-May-2015.

[47] K. Xu, D. Feng, and H. Mi, "Deep convolutional neural network-based early

automated detection of diabetic retinopathy using fundus image," *Molecules*, vol. 22, no. 12, Dec. 2017.

[48] H. Greenspan, B. Van Ginneken, and R. M. Summers, "Guest Editorial Deep Learning in Medical Imaging: Overview and Future Promise of an Exciting New Technique," *IEEE Transactions on Medical Imaging*, vol. 35, no. 5. Institute of Electrical and Electronics Engineers Inc., pp. 1153–1159, 01-May-2016.

[49] M. Matsugu, K. Mori, Y. Mitari, and Y. Kaneda, "Subject independent facial expression recognition with robust face detection using a convolutional neural network," in *Neural Networks*, 2003, vol. 16, no. 5–6, pp. 555–559.

[50] O. K. Oyedotun, E. O. Olaniyi, and A. Khashman, "Disk hernia and spondylolisthesis diagnosis using biomechanical features and neural network," *Technol. Heal. Care*, vol. 24, no. 2, pp. 267–279, Mar. 2016.

[51] D. Shen, G. Wu, and H.-I. Suk, "Deep Learning in Medical Image Analysis," *Annu. Rev. Biomed. Eng.*, vol. 19, no. 1, pp. 221–248, Mar. 2017.

[52] S. Indolia, A. K. Goswami, S. P. Mishra, and P. Asopa, "Conceptual Understanding of Convolutional Neural Network- A Deep Learning Approach," in *Procedia Computer Science*, 2018, vol. 132, pp. 679–688.

[53] K. O'Shea and R. Nash, "An Introduction to Convolutional Neural Networks," Nov. 2015.

[54] A. Voulodimos, N. Doulamis, A. Doulamis, and E. Protopapadakis, "Deep Learning for Computer Vision: A Brief Review," *Computational Intelligence and Neuroscience*, vol. 2018. Hindawi Limited, 2018.

[55] A. L. Maas, A. Y. Hannun, and A. Y. Ng, "Rectifier Nonlinearities Improve Neural Network Acoustic Models," 2013.

[56] K. He, X. Zhang, S. Ren, and J. Sun, "Delving Deep into Rectifiers: Surpassing Human-Level Performance on ImageNet Classification."

[57] N. Srivastava, G. Hinton, A. Krizhevsky, and R. Salakhutdinov, "Dropout: A Simple Way to Prevent Neural Networks from Overfitting," 2014.

[58] Y. Bengio, "Practical Recommendations for Gradient-Based Training of Deep Architectures BT - Neural Networks: Tricks of the Trade: Second Edition," G. Montavon, G. B. Orr, and K.-R. Müller, Eds. Springer Berlin Heidelberg, 2012, pp. 437–478.

[59] J. Olczak *et al.*, "Artificial intelligence for analyzing orthopedic trauma

radiographs.," *Acta Orthop.*, vol. 88, no. 6, pp. 581–586, Dec. 2017.

[60] A. Prasoon, K. Petersen, C. Igel, F. Lauze, E. Dam, and M. Nielsen, "Deep feature learning for knee cartilage segmentation using a triplanar convolutional neural network," in *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 2013, vol. 8150 LNCS, no. PART 2, pp. 246–253.

[61] H. Lee *et al.*, "Fully Automated Deep Learning System for Bone Age Assessment," *J. Digit. Imaging*, vol. 30, no. 4, pp. 427–441, Aug. 2017.

[62] K. J. Dreyer and J. R. Geis, "When Machines Think: Radiology's Next Frontier," *Radiology*, vol. 285, no. 3, pp. 713–718, Nov. 2017.

[63] B. Q. Huynh, H. Li, and M. L. Giger, "Digital mammographic tumor classification using transfer learning from deep convolutional neural networks," *J. Med. Imaging*, vol. 3, no. 3, p. 034501, Aug. 2016.

[64] I. Banerjee, A. Crawley, M. Bhethanabotla, H. E. Daldrup-Link, and D. L. Rubin, "Transfer learning on fused multiparametric MR images for classifying histopathological subtypes of rhabdomyosarcoma.," *Comput. Med. Imaging Graph.*, vol. 65, pp. 167–175, 2018.

[65] "Computer-aided diagnosis in medical imaging: Historical review, current status and future potential," *Comput. Med. Imaging Graph.*, vol. 31, no. 4–5, pp. 198–211, Jun. 2007.

[66] S. Singh, J. Maxwell, J. A. Baker, J. L. Nicholas, and J. Y. Lo, "Computer-aided Classification of Breast Masses: Performance and Interobserver Variability of Expert Radiologists versus Residents," *Radiology*, vol. 258, no. 1, pp. 73–80, Jan. 2011.

[67] J. Z. Cheng *et al.*, "Computer-Aided Diagnosis with Deep Learning Architecture: Applications to Breast Lesions in US Images and Pulmonary Nodules in CT Scans," *Sci. Rep.*, vol. 6, Apr. 2016.

[68] S. Pereira, A. Pinto, V. Alves, and C. A. Silva, "Brain Tumor Segmentation Using Convolutional Neural Networks in MRI Images," *IEEE Trans. Med. Imaging*, vol. 35, no. 5, pp. 1240–1251, May 2016.

[69] A. A. A. Setio *et al.*, "Pulmonary Nodule Detection in CT Images: False Positive Reduction Using Multi-View Convolutional Networks.," *IEEE Trans. Med. Imaging*, vol. 35, no. 5, pp. 1160–1169, 2016.

[70] H. Sharma, N. Zerbe, I. Klempert, O. Hellwich, and P. Hufnagl, "Deep

convolutional neural networks for automatic classification of gastric carcinoma using whole slide images in digital histopathology," *Comput. Med. Imaging Graph.*, vol. 61, pp. 2–13, Nov. 2017.

[71] M. Cicero *et al.*, "Training and Validating a Deep Convolutional Neural Network for Computer-Aided Detection and Classification of Abnormalities on Frontal Chest Radiographs," *Invest. Radiol.*, vol. 52, no. 5, pp. 281–287, May 2017.

[72] J. Duryea, S. Zaim, and F. Wolfe, "Neural network based automated algorithm to identify joint locations on hand/wrist radiographs for arthritis assessment," *Med. Phys.*, vol. 29, no. 3, pp. 403–411, 2002.

[73] S. Murakami, K. Hatano, J. Tan, H. Kim, and T. Aoki, "Automatic identification of bone erosions in rheumatoid arthritis from hand radiographs based on deep convolutional neural network," *Multimed. Tools Appl.*, vol. 77, no. 9, pp. 10921–10937, 2018.

[74] M. B.-H.-… C. de Física and undefined 2018, "Automatic Diagnosis of Rheumatoid Arthritis from Hand Radiographs using Convolutional Neural Networks," *revistacubanadefisica.org*.

[75] C. Spampinato, S. Palazzo, D. Giordano, M. Aldinucci, and R. Leonardi, "Deep learning for automated skeletal bone age assessment in X-ray images," *Med. Image Anal.*, vol. 36, pp. 41–51, Feb. 2017.

[76] H. Shin, H. Roth, M. Gao, L. Lu, … Z. X.-I. transactions on, and undefined 2016, "Deep convolutional neural networks for computer-aided detection: CNN architectures, dataset characteristics and transfer learning," *ieeexplore.ieee.org*.

[77] Y. Mednikov, S. Nehemia, B. Zheng, O. Benzaquen, and D. Lederman, "Transfer Representation Learning using Inception-V3 for the Detection of Masses in Mammography," in *Proceedings of the Annual International Conference of the IEEE Engineering in Medicine and Biology Society, EMBS*, 2018, vol. 2018-July, pp. 2587–2590.

[78] D. Kim, T. M.-C. radiology, and undefined 2018, "Artificial intelligence in fracture detection: transfer learning from deep convolutional neural networks," *Elsevier*.

[79] Y. Xue, R. Zhang, Y. Deng, K. Chen, and T. Jiang, "A preliminary examination of the diagnostic value of deep learning in hip osteoarthritis,"

PLoS One, vol. 12, no. 6, Jun. 2017.

[80] A. Tiulpin, J. Thevenot, E. Rahtu, P. Lehenkari, and S. Saarakkala, "Automatic knee osteoarthritis diagnosis from plain radiographs: A deep learning-based approach," *Sci. Rep.*, vol. 8, no. 1, Dec. 2018.