Chapter 2
Literature Review

A literature review on machine learning for pneumonia detection usually includessummarizing and analyzing the pertinent literature, reports, and academic papers. It seeks to give a broad overview of the state of the research, point out research gaps, and highlight new trends and directions. Here is a summary of the main topics that are frequently discussed in a literature review on machine learning may be introduction and background, problem formulation and application, types of machine learning algorithms, Evaluation Metrics and Performance Measures, Feature Selection and Engineering, Model Training and Optimization, Ethical and Fairness Considerations, Emerging Trends and Future Directionsetc

2.1 Pneumonia and its Diagnostic Challenges

This chapter focuses on pneumonia, a common respiratory infection that can be fatal, as well as the difficulties in diagnosing it. Due to the substantial morbidity and mortality it causes worldwide, pneumonia is an important global health concern. However, because of a number of variables, precisely diagnosing pneumonia can be challenging.

2.1.1 Definition and Overview of Pneumonia

An extensive explanation of pneumonia is given in this section. It describes the infection's characteristics, its causes, and the various forms of pneumonia, such as aspiration pneumonia, community-acquired pneumonia, and pneumonia obtained in a hospital. Additionally, the section emphasizes how crucial early and precise diagnosis is for successful management and therapy.

A respiratory infection called pneumonia is characterized by inflammation of the lung's tissue, particularly the alveoli (air sacs). Bacteria, viruses, fungus, or other microbes are frequently to blame for it. All ages can be affected by pneumonia, but it typically worsens in the elderly, small children, and people with compromised immune systems.

Inhaling infectious droplets, aspirating foreign objects into the lungs, or bacteria spreading via the bloodstream from other regions of the body are only a few causes of pneumonia. The infection causes the lungs to become inflamed, which blocks the normal exchange of oxygen and carbon dioxide by causing the alveoli to fill with fluid and inflammatory cells.

Depending on the aetiology, patient age, and underlying medical conditions, pneumonia can manifest clinically in a variety of ways. Fever, cough (with or without sputum production), exhaustion, shortness of breath, chest pain, and occasionally disorientation or changed mental status in older persons are common symptoms. In severe circumstances, complications from pneumonia can include organ damage, sepsis, and respiratory failure.

Based on the environment in which the infection occurred, there are many forms of pneumonia. Pneumonia cases that are "community-acquired" (CAP) are those that occur outside of a hospital setting, like in the neighborhood or at home. Hospital-acquired pneumonia (HAP) happens while a patient is in the hospital, usually 48 hours after admission. Ventilator-associated pneumonia (VAP), which manifests in patients undergoing mechanical ventilation in an intensive care unit, is a different form.

Clinical assessment, radiographic imaging, and microbiological testing are all used to diagnose pneumonia. Healthcare professionals review risk factors, conduct a physical exam, and evaluate the patient's symptoms. The lungs can be seen on chest X-rays and computed tomography (CT) scans, which can also be utilised to spot distinctive abnormalities such infiltrates or consolidation. The exact causing pathogen can be identified with the aid of microbiological testing, such as sputum cultures, blood cultures, and molecular tests.

But because the symptoms of pneumonia and other respiratory illnesses overlap, making a diagnosis of pneumonia difficult. Recent developments in biomarkers, quick diagnostic tests, and emerging technologies like machine learning and artificial intelligence have showed promise in enhancing the precision and efficacy of pneumonia diagnosis.

In conclusion, pneumonia is an inflammatory lungs illness brought on by a variety of microbes. It manifests as respiratory symptoms and, especially in vulnerable groups, can result in serious consequences. Clinical evaluation, radiographic imaging, and microbiological testing are used in the diagnosis, albeit it might be difficult to pinpoint the exact cause. The goal of ongoing research and technological development is to improve pneumonia diagnosis and patient outcomes.

2.1.2 Clinical Presentation and Symptoms

This section focuses on the clinical presentation and symptoms of pneumonia. It discusses the typical signs and symptoms associated with the infection, such as fever, cough, chest pain, and shortness of breath. The section also explores the variations in symptoms based on age, underlying health conditions, and the specific causative agents involved.

2.1.3 Radiological Imaging for Pneumonia Diagnosis

Radiological imaging plays a crucial role in diagnosing pneumonia. This section explores the different imaging modalities used in pneumonia diagnosis, including chest X-rays and computed tomography (CT) scans. It discusses the radiographic findings associated with pneumonia, such as consolidation, infiltrates, and pleural effusion. The section also addresses the limitations and challenges of relying solely on imaging for accurate diagnosis.

Pneumonia is an infection that can be brought on by a virus, bacteria, fungus, or other germs and results in inflammation in one or both of the lungs. To diagnose the patient's condition, the doctor may perform a physical examination in addition to using a chest x-ray, chest CT, chest ultrasound, or lung biopsy by needle. With the aid of thoracentesis, chest tube insertion, or image-guided abscess drainage, the doctor may further assess the patient's health and lung function.

An illness called pneumonia can result in inflammation in one or both lungs. A virus, bacteria, fungi, or other types of germs may be to blame. Typically, a person contracts the sickness by inhaling air contaminated with microorganisms.

The following signs and symptoms could be present in pneumonia patients:

- Cough that produces phlegm or sometimes blood
- Fever
- Shortness of breath or difficulty breathing
- Chills or shaking
- Fatigue
- Sweating
- Chest or muscle pain

People with existing health problems are also at increased risk. Risk factors and circumstances that may increase a person's chances of developing pneumonia include:

- Having illnesses such as emphysema, HIV/AIDS or other lung diseases or conditions that affect the immune system
- Having the flu
- Exposure to and inhalation of various chemicals
- Smoking or excessive drinking
- A prolonged stay in the hospital or intensive care
- Recent surgery
- Recent injury

Pneumonia can sometimes lead to serious complications, such as respiratory system failure, spread of infections, fluid surrounding the lungs, abscesses or uncontrolled inflammation throughout the body (sepsis). The condition can also be fatal, so it is important to seek immediate medical attention if the patientis experiencing these symptoms. Primary doctor will begin by asking the patient about his/her medical history and symptoms. The Patient will also undergo a physical exam, so that the doctor can listen to patient’s lungs. In checking for pneumonia, the doctor will listen for abnormal sounds like crackling, rumbling or wheezing. If the doctor thinks the patient may have pneumonia, an imaging test may be performed to confirm the diagnosis.

To check for pneumonia, one or more of the following tests might be prescribed:

- **Chest x-ray:** During an x-ray examination, your doctor will be able to see your heart, blood vessels, and lungs to help determine whether you have pneumonia. The radiologist will search for infiltrates, which are white areas in the lungs that signify an infection, when interpreting the x-ray. This examination will also assist in identifying any pneumonia-related problems, such as abscesses or pleural effusions (fluid surrounding the lungs), that you may have.

(a) Normal (b) Pneumonia (c) COVD-19

Figure2.1: X-Ray of Normal, Pneumonia, COVID-19 Patient [11]

- **CT of the lungs:** A CT scan of the chest can be performed to detect pneumonia that may be more difficult to spot on a normal x-ray and to see finer features within the lungs. A CT scan can assist establish whether pneumonia may be due to an issue with the airway by providing a very detailed image of the airway (trachea and bronchi). A CT scan can also detect enlarged lymph nodes, abscesses, pleural effusions, and complications from pneumonia.

Figure2.2: Chest CT Findings of Patients Infected With Novel CoronavirusPneumonia [11]

- **Ultrasound of the chest:**If fluid around the lungs is suspected, an ultrasound of the chest may be used. The amount of fluid present can be assessed using an ultrasound, which can also help identify the source of the fluid.
- **MRI of the chest:** Although MRI is occasionally used to examine the heart, chest arteries, and components of the chest wall, it is not typically utilised to assess for pneumonia. An MRI may reveal more details regarding the origin or severity of lungs abnormalities caused by an infection, tumour, or excess fluid.
- **Needle biopsy of the lung:** To identify the cause of pneumonia, your doctor might ask for a lung biopsy. This process entails taking a number of little samples from your lung or lung(s) and analysing them. Lung biopsies can be performed with the use of an MRI, CT, ultrasound, or x-ray.

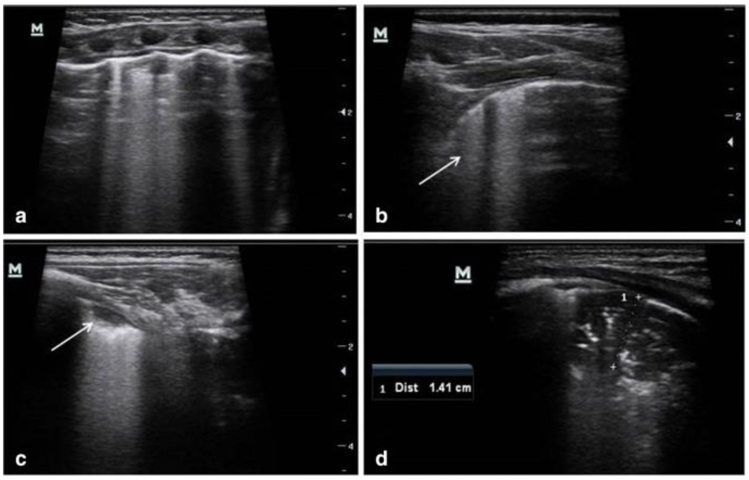

Figure2.3: Lung ultrasound for the diagnosis of pneumonia in children [11]

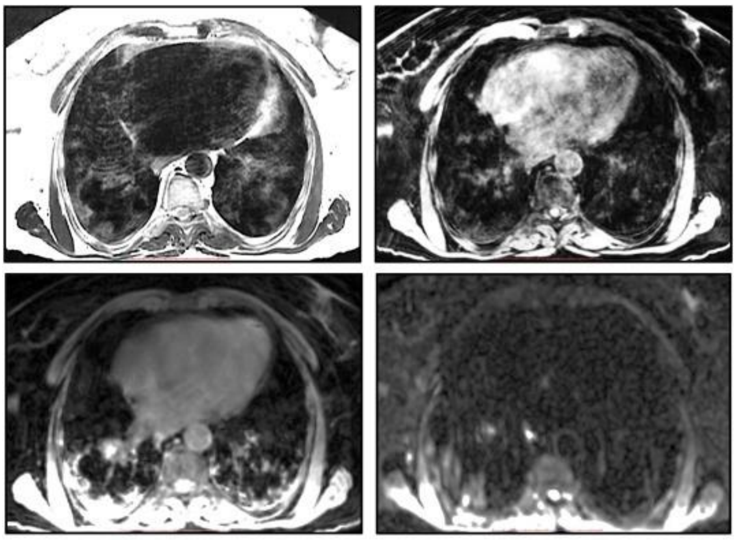

Figure2.4: Chest MRI of patients with COVID-19 [11]

2.1.4 Microbiological Testing for Pneumonia

Microbiological testing is essential for identifying the causative agent in pneumonia. This section discusses the various methods used for microbiological testing, including sputum cultures, blood cultures, and polymerase chain reaction (PCR) tests. It explores the advantages and limitations of each method, including issues of sample quality, false-negative results, and the time required for obtaining results.

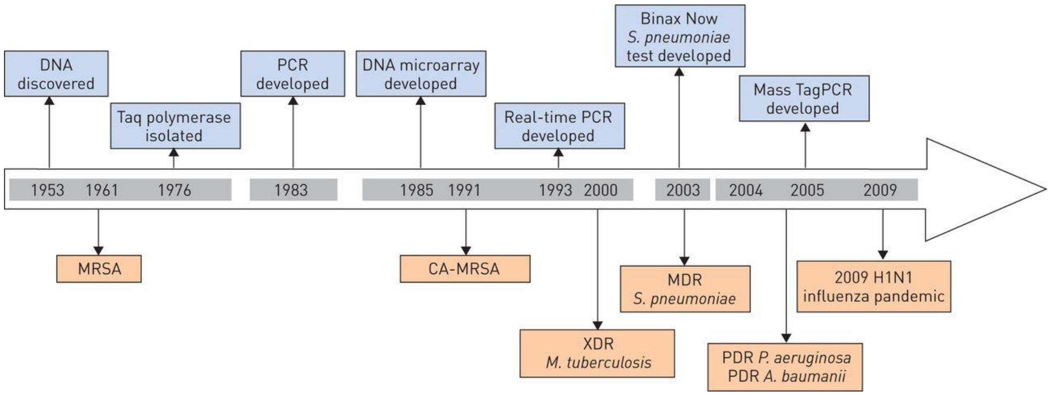

Figure2.5: Laboratory diagnosis of pneumonia in the molecular age [12]

2.1.5 Challenges in Pneumonia Diagnosis

This section highlights the diagnostic challenges faced in accurately diagnosing pneumonia. It discusses the limitations of clinical assessment, as symptoms can overlap with other respiratory conditions. The section also addresses the challenges associated with radiological imaging, including the interpretation of findings and the potential for false-positive or false-negative results. Furthermore, it explores the limitations of microbiological testing, such as the difficulty in obtaining appropriate samples and the time required for test results.

2.1.6 Advances in Pneumonia Diagnostics

This section explores the recent advancements in pneumonia diagnostics. It discusses the development of rapid diagnostic tests, such as point-of-care antigen tests and molecular assays, which can provide quicker and more accurate results. The section also highlights the role of biomarkers and emerging technologies, such as machine learning and artificial intelligence, in improving pneumonia diagnosis.

2.2 Medical Imaging in Pneumonia Diagnosis & Other applications for image analysis

The diagnosis and treatment of many medical disorders, including pneumonia, depend heavily on medical imaging. In order to diagnose pneumonia and for other image analysis applications, medical imaging is employed as follows:

2.2.1Pneumonia Diagnosis

Pneumonia is frequently diagnosed through imaging methods including computed tomography (CT) scans and chest X-rays. These imaging techniques aid in the visualization of the lungs and the detection of areas with pneumonia-specific swelling, consolidation, or fluid collection. To confirm the presence of pneumonia, assess its severity and extent, and inform treatment choices, radiologists examine these images.

2.2.2Treatment Monitoring

Monitoring the effectiveness of the treatment for pneumonia also uses medical imaging. Follow-up chest X-rays or CT scans can be used to determine whether lung infiltrates have cleared up or whether complications like abscesses or pleural effusion are present. These imaging tests help determine the effectiveness of the treatment and provide direction for subsequent care.

2.2.3Other Respiratory Conditions

Medical imaging is helpful in identifying and treating other respiratory disorders in addition to pneumonia. For instance, lung tumours, chronic obstructive pulmonary disease (COPD), interstitial lung disorders, and pulmonary embolism, among other conditions, are all assessed using chest X-rays and CT scans. These imaging methods offer thorough anatomical details that help with precise diagnosis and therapy planning.

2.2.4Cardiovascular Imaging

Medical imaging techniques like echocardiography, cardiac computed tomography angiography (CCTA), and cardiac magnetic resonance imaging (MRI) are used to assess heart structure and function. These imaging modalities are essential in diagnosing and managing conditions such as coronary artery disease, heart failure, valvular diseases, and congenital heart abnormalities.

2.2.5Neuroimaging

The brain and spinal cord can be seen using neuroimaging methods including computed tomography (CT) and magnetic resonance imaging (MRI). For the diagnosis and follow-up of a number of neurological disorders, such as stroke, brain tumours, multiple sclerosis, and traumatic brain injuries, these imaging modalities are crucial.

2.2.6Oncology

In oncology, medical imaging is essential for cancer diagnosis, staging, and therapy planning. Imaging techniques like CT, MRI, PET, and mammography help visualisetumours, spot metastases, evaluate therapy effectiveness, and track the course of the disease.

2.2.7Image-Guided Interventions

Medical imaging is used to guide minimally invasive procedures, such as image-guided biopsies, radiofrequency ablation, and catheter-based interventions. Real-time imaging, such as ultrasound, fluoroscopy, and CT, helps visualize the target area, guide the needle or catheter, and ensure accurate placement and treatment delivery.

2.2.8Musculoskeletal Imaging

Radiography, CT, MRI, and ultrasound are commonly used in musculoskeletal imaging to assess bone fractures, joint disorders, soft tissue injuries, and tumors. These imaging modalities aid in diagnosis, surgical planning, and monitoring treatment response in orthopedic and rheumatologic conditions.

These are only a handful of the numerous uses for medical imaging and image analysis in different medical specialties. Emerging technologies like artificial intelligence (AI) and machine learning are being used into medical imaging to improve efficiency, increase diagnostic accuracy, and enable more individualized patient care.

2.3 Deep Learning for Medical Image Analysis

Deep learning has completely changed the way medical image analysis is done by making it possible to automatically and accurately interprets medical pictures. Some essential elements of deep learning in medical image analysis are listed below:

Convolutional Neural Networks (CNNs): The cornerstone of deep learning in medical image processing is convolutional neural networks. Through numerous layers of convolution, pooling, and non-linear activations, CNNs excel at extracting significant information from images. These networks are made to recognise patterns from the pixel values of medical images and learn how to classify them.

Segmentation: In medical pictures, deep learning models can precisely segment anatomical features or abnormalities. Delineating regions of interest within an image, such as tumours, organs, or blood arteries, is known as segmentation. Many different medical sectors have successfully used CNN-based architectures for segmentation tasks, such as U-Net, SegNet, or Mask R-CNN.

Classification and Diagnosis: Exceptional performance in image-based categorization and diagnosis has been shown by deep learning models. CNNs can categorise medical images into several illness categories by learning discriminative features through training on big datasets. For instance, CNNs have been used to identify Alzheimer's disease from brain MRI scans, diabetic retinopathy from fundus pictures, and lung cancer from CT scans.

Object Detection: Deep learning models have the ability to locate and identify particular things in medical photos. For instance, they can pinpoint the presence and location of abnormalities including lung nodules, breast tumours, and brain haemorrhages in radiology. For use in medical image analysis applications, object detection techniques as Faster R-CNN, YOLO, or SSD have been modified.

Generative Adversarial Networks (GANs): A deep learning model called a GAN can produce synthetic medical images with accurate characteristics. In order to address concerns with data scarcity, GANs have been used for data augmentation, where they can produce more training samples. In order to increase picture registration or improve image quality, they have also been used for image-to-image translation tasks, such as transforming CT scans into MRI-like images.

Transfer Learning: Transfer learning allows for the adaptation of deep learning models for use in medical image analysis from broad image datasets like ImageNet. Smaller medical picture datasets can be used to fine-tune pre-trained CNNs, expediting model training and enhancing performance.

Radiomics and Feature Extraction: To enable radiomic analysis, deep learning models can be used to extract high-dimensional features from medical images. The process of radiomics entails removing quantitative elements from images that represent traits like shape, texture, or intensity changes. These characteristics can also be utilised to forecast therapy outcomes, illness development, or patient survival.

Explainability and Interpretability: Deep learning model interpretation in medical image processing is a current research topic. Techniques like Grad-CAM, saliency maps, and attention mechanisms can be used to emphasise key areas in an image and offer justifications for model predictions. Forging trust and promoting clinical adoption, model interpretability is essential.

Medical image analysis has substantially advanced thanks to deep learning, and this advancement has the potential to increase diagnostic precision, help with treatment planning, and improve patient care. To realize these models' full potential in actual healthcare settings, it is crucial to provide thorough validation, ethical considerations, and integration with clinical workflows.

2.4 Convolutional Neural Networks (CNNs)

CNN are a class of deep learning models specifically designed for processing and analyzing grid-like structured data, such as images. They have revolutionized computer vision tasks and achieved state-of-the-art performance in various image-related applications. Here are some key aspects of CNNs:

Convolutional Layers: CNNs are built upon convolutional layers that perform local receptive field operations. Convolutional filters are applied to small regions of the input image, computing element-wise multiplications and summing the results to produce a feature map. These filters capture local spatial patterns, such as edges, corners, or textures, and enable hierarchical feature extraction.

Pooling Layers: Pooling layers are typically inserted after convolutional layers to downsample the feature maps and reduce spatial dimensions. Max pooling is a common pooling operation that retains the maximum value within each pooling region, effectively reducing the spatial resolution while preserving the most salient features. Pooling helps achieve translation invariance and reduces computational complexity.

Non-Linear Activation Functions: Activation functions introduce non-linearity into CNNs, enabling them to model complex relationships in the data. Rectified Linear Units (ReLU) activation function is commonly used in CNNs, as it replaces negative values with zero and keeps positive values unchanged. ReLU promotes faster training convergence and avoids the vanishing gradient problem.

Fully Connected Layers: After several convolutional and pooling layers, CNNs are often followed by one or more fully connected layers. These layers connect all the neurons from the previous layer to the current layer, performing high-level feature integration and classification. Fully connected layers are typically implemented using densely connected neural network architectures.

Training with Backpropagation: CNNs are trained using the backpropagation algorithm, where the model's weights are updated based on the gradients of the loss function with respect to the model's predictions. The gradients are propagated backward through the layers, allowing the model to learn and adjust the filters' parameters to minimize the prediction errors.

Transfer Learning: CNNs trained on large-scale image datasets, such as ImageNet, can be used as a starting point for other image analysis tasks through transfer learning. Pretrained CNNs are fine-tuned on smaller task-specific datasets, accelerating the training process and leveraging the learned general image representations.

Data Augmentation: Data augmentation techniques are commonly employed with CNNs to artificially increase the diversity of the training data. By applying random transformations such as rotations, translations, flips, or distortions, the CNN is exposed to a wider range of variations and becomes more robust to changes in the input data.

Architectural Variations: CNN architectures can vary in complexity and depth, depending on the task and dataset size. Popular CNN architectures include AlexNet, VGGNet, GoogLeNet (Inception), ResNet, and DenseNet, among others. These architectures differ in the number of layers, filter sizes, skip connections, and other architectural choices, aiming to improve performance and address challenges like vanishing gradients or over fitting.

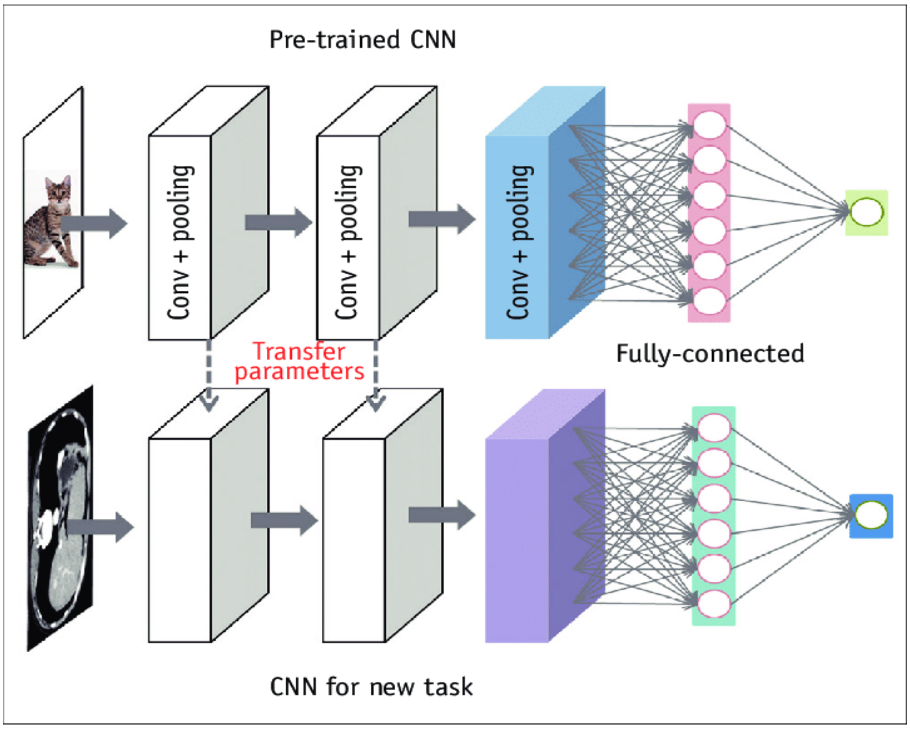

Figure2.6: Transfer Learning Process [8]

CNNs have achieved remarkable success in numerous computer vision tasks, including image classification, object detection, image segmentation, and more. Their ability to automatically learn and extract meaningful features from images, combined with their hierarchical structure, has made CNNs a powerful tool in analyzing and understanding visual data.

2.5 Multimodal Learning Approaches

Multimodal learning approaches involve combining information from multiple modalities, such as text, images, audio, or sensor data, to improve learning and decision-making in machine learning models. Here are some key aspects of multimodal learning approaches:

Fusion Techniques: Multimodal learning requires integrating information from different modalities effectively. Fusion techniques combine the features or representations extracted from each modality to create a unified representation that captures the complementary information. Fusion can be performed at different levels, such as early fusion (combining features at the input level), late fusion (combining outputs of individual modality models), or intermediate fusion (combining representations at intermediate layers).

Deep Multimodal Architectures: Deep learning models, such as multimodal deep neural networks (DNNs), have been developed to effectively handle multimodal data. These architectures typically consist of multiple pathways, each dedicated to processing a specific modality. The pathways can share lower-level layers to capture shared representations, while higher-level layers are modality-specific. The models are trained jointly to learn multimodal representations and make predictions based on the combined information.

Cross-Modal Embeddings: Cross-modal embeddings aim to project data from different modalities into a common latent space, where the information from different modalities is aligned. These embeddings facilitate cross-modal retrieval, matching, or clustering tasks. Techniques like Canonical Correlation Analysis (CCA), Deep

Canonical Correlation Analysis (DCCA), or Joint Embedding models aim to learn shared representations that capture the underlying relationships between modalities.

Attention Mechanisms: Attention mechanisms have proven valuable in multimodal learning by allowing models to focus on the most relevant information from each modality. These mechanisms assign weights to different modalities or parts of modalities based on their relevance to the task. Attention-based multimodal architectures enable the model to selectively attend to important cues and improve performance.

Modality Alignment and Alignment Losses: Multimodal learning often requires aligning the representations from different modalities to ensure compatibility. Alignment techniques aim to learn a shared representation space by minimizing the discrepancy between modalities. Alignment losses, such as contrastive losses or triplet losses, encourage similar representations for related instances from different modalities while pushing apart unrelated instances.

Data Augmentation and Modality Dropout: Data augmentation techniques can be applied independently to each modality to increase the diversity and variability of the training data. Modality dropout is another technique where individual modalities are randomly dropped during training to encourage the model to rely on the remaining modalities, improving robustness and generalization.

Transfer Learning and Pre-training: Pre-training multimodal models on large-scale datasets or pre-training individual modality-specific models can be beneficial for multimodal learning tasks. Pre-training can provide useful initializations and leverage the learned representations to improve performance on specific multimodal tasks.

Applications: Multimodal learning approaches find applications in various domains. Some examples include multimodal sentiment analysis, audio-visual speech recognition, visual question answering, human activity recognition using sensor data, or healthcare applications combining medical imaging and clinical data.

Multimodal learning approaches enable models to leverage information from different modalities, capturing rich and complementary information that enhances learning avnd decision-making. These techniques are valuable in tasks where multiple modalities contribute to the understanding and interpretation of data, leading to more robust and accurate models**.**

2.6 Transfer Learning in Medical Image Analysis

Transfer learning in medical image analysis involves leveraging pretraining on large-scale general image datasets and applying the learned knowledge to improve the performance of deep learning models on medical imaging tasks. Here's how transfer learning is applied in medical image analysis:

Pretrained Models: Deep learning models pretrained on general image datasets, such as ImageNet, have learned rich representations of visual features that can be transferred to medical image analysis tasks. These pretrained models, such as convolutional neural networks (CNNs) like VGGNet, ResNet, or InceptionNet, can be used as a starting point for medical image analysis.

Feature Extraction: In transfer learning, the pretrained models serve as powerful feature extractors. The initial layers of the pretrained models capture low-level visual features like edges, textures, or shapes, which are transferable across different domains. These pretrained models can be used to extract features from medical images, and these features are then fed into task-specific classifiers or other downstream models.

Fine-tuning: After extracting features from the pretrained models, the transferred features can be used as input to train additional layers or classifiers specifically for the medical image analysis task. Fine-tuning involves adjusting the parameters of the pretrained model, typically the higher-level layers, to better adapt to the specific medical domain. This process helps the model specialize in the medical image analysis task while leveraging the general visual representations learned from the pretrained model.

Data Augmentation: Data augmentation techniques are commonly employed in transfer learning for medical image analysis. By applying random transformations to the medical images, such as rotations, translations, or flips, the dataset is augmented with additional variations. Data augmentation helps improve the generalization of the model by exposing it to a wider range of image variations and reducing overfitting.

Limited Labeled Data: Transfer learning is particularly useful when there is a scarcity of labeled medical image data. Pretrained models trained on general image datasets have been exposed to vast amounts of labeled data, allowing them to learn robust visual representations. By utilizing these pretrained models, the need for large amounts of labeled medical image data is mitigated, and the model can still benefit from the knowledge captured by the pretrained models.

Domain Adaptation: In medical image analysis, the domain shift between the general image dataset and the medical images can present a challenge. Domain adaptation techniques aim to reduce the distribution discrepancy between the source (general image dataset) and target (medical image dataset) domains. Adapting the pretrained models to the medical domain can involve various strategies, such as domain adversarial training, feature alignment, or domain-specific fine-tuning.

Transfer learning has been successfully applied in various medical image analysis tasks, including image classification, segmentation, detection, and disease diagnosis. It allows models to leverage the knowledge learned from large-scale general image datasets and significantly improves performance, especially in scenarios with limited labeled medical image data.

A model that has been trained on one job can be used as a starting point for developing a model on another, related task using the machine learning technique known as transfer learning. Transfer learning uses the information and learned representations from the pre-trained model to speed up the learning process and enhance performance on the target task rather than beginning the training process from scratch.Research article and research findings related to the artificial intelligence use in image analysis is described in next section.

2.7 Summary &pre research findings

Research Paper: 001

Title :	"Face Recognition: A Convolutional Neural-Network Approach"
Publication :	IEEE TRANSACTIONS ON NEURAL NETWORKS, VOL. 8, NO. 1, JANUARY 1997
Authors:	Steve Lawrence, Member, IEEE, C. Lee Giles, Senior Member, IEEE, Ah Chung Tsoi, Senior Member, IEEE,and Andrew D. Back, Member, IEEE

<u>LEARNING</u>

Methods Used	• KarhunenLoeve (KL) instead of Self Organizing Map (SOM-CNN) • Multilayer Perceptron (MLP) instead of Convolutional Neural Network
Merits/ Demerits/ Outcomes	Merit As well (5.3% error versus 3.8%) Demerit Poor (40% error versus 3.8%)
Possible Improvements	• More careful selection of convolutional architecture can increase the recognition efficiency • The various facial features could be ranked according to their importance in recognizing faces and separate modules could be introduced for various parts of the face, e.g., the eye region, the nose region, and the mouth region obtain very good performance using a simple template matching strategy on precisely these regions)

Research Paper: 002

Title :	"Image classification with an RGB-channel nonsubsampledcontourlet transform and a convolutional neural network"
Publication :	Neurocomputing, Elsevier (2019)

Authors:	Lingling Fang*, Hanyu Zhang, Jiaxin Zhou, Xianghai Wang Department of Computing and Information Technology, Liaoning Normal University, Dalian City, China

LEARNING

Methods Used	Nonsubsampledcontourlet transform (NSCT) for image classification Then Convolutional Neural Network for detection.
Merits/ Demerits/ Outcomes	Merits Classifies color images without incorporating any prior knowledge. Represents the statistical properties and reduce the information redundancy of the RGB channel natural image
Possible Improvements	Different classification methods combined with CNN

Research Paper: 003

Title :	"Super-resolution Reconstruction of Single Anisotropic 3D MR Images Using Residual Convolutional Neural Network"
Publication :	Neurocomputing, Elsevier (2019)
Authors:	Jinglong Du, Zhongshi He, Lulu Wang, Ali Gholipour, Zexun Zhou, Dingding Chen, YuanyuanJi College of Computer Science, Chongqing University, Chongqing 400044, China

LEARNING

Methods Used	A novel CNN-based anisotropic MR image reconstruction method
Merits/ Demerits/ Outcomes	Merit Over performs interpolation methods, non-local mean method (NLM) and sparse coding based algorithm in terms of peak SNR,

	structural similarity index, intensity profile and small structures. Demerit hard-ware limitations and time constraints Result in the acquisition of anisotropic MR images
Possible Improvements	Developing new algorithm of convolutional neural network by changing the mapping sequence and neural network parameters like no. of nodes, mapping algorithm etc Improvement like less computational time, less memory usage.

Research Paper: 004

Title :	"Brain Tumor Segmentation Based on Improved Convolutional Neural Network in Combination with Non-quantifiable LocalTexture Feature"
Publication :	Journel of Medical Systems – Image & Signal Processing Springer (2019)
Authors:	JWu Deng, Qinke Shi, Kai Luo, Yi Yang, NingNing Information Center, West China Hospital of Sichuan university, Chengdu 610000, Sichuan, China

LEARNING

Methods Used	FCNN & Dense Micro Block Difference Feature (DMDF)
Merits/ Demerits/ Outcomes	Merits Compared to traditional MRI brain tumor methods, the experimental results show the significant accuracy and stability improved. Very high real-time performance as brain tumor image can segment within 1 s.
Possible Improvements	Comparison with FCNN, IFCNN, Pereiral, Havaei More methods can be compared with the new proposed method.

Research Paper: 005

Title :	“Cucumber leaf disease identification with global pooling dilated convolutional neural network”
Publication :	Computer & Electronics in Agriculture, Elsevier-Science Direct (2019)
Authors:	Shanwen Zhang, Subing Zhang, Chuanlei Zhang, Xianfeng Wang, Yun Shi College of Computer Science and Information Engg., Tianjin University of Science and Technology, Tianjin 300222, China

LEARNING

Methods Used	Global pooling dilated convolutional neural network (GPDCNN) combining dilated convolution with global pooling
Merits/ Demerits/ Outcomes	Merits • Convolution receptive field are increased without increasing complexity compared to traditional CNN • Dilated convolutional layer is employed to recover the spatial resolution without increasing the number of training parameters • Getting Merits of both Global pooling & dilated convolution
Possible Improvements	• With dilated convolutional layers, GPDCNN can extend the receptive field without losing resolution • Improving the GPDCNN performance by exploring the key role of probabilistic graphical models • Improving the GPDCNN that can find the crop dieses or Crop Dieses Recognition System (CDRS)

Research Paper: 006

Title :	"Large-scale Video Classification with Convolutional Neural Networks"
Publication :	IEEE Xplore, Computer Vision Foundation (2014)
Authors:	Andrej Karpathy, George Toderici, SankethShetty, Thomas Leung, Rahul Sukthankar, Li Fei-Fei, Google Research, Computer Science Department, Stanford University

LEARNING

Methods Used	Convolutional Neural Network for video classification
Merits/ Demerits/ Outcomes	1 million YouTube videos belonging to 487 classes Merits Spatio-temporal networks display significant performance improvement compared to strong feature-based base lines (55.3% to 63.9%).
Possible Improvements	• Single frame detection will be more efficient then prediction with multi-frame • They have experimented sports videos, we can research on another categories • Exploring recurrent neural networks as a more powerful technique for combining clip-level predictions into global video-level predictions

Research Paper: 007

Title :	"A Convolutional Neural Network Cascade for Face Detection"
Publication :	IEEE Xplore, Computer Vision Foundation (2015)

Authors:	Haoxiang Li, Zhe Lin, XiaohuiShen, Jonathan Brandt, Gang Hua Adobe Research, San Jose, California

LEARNING

Methods Used	cascade architecture built on convolutional neural networks (CNNs)
Merits/ Demerits/ Outcomes	Introduced a CNN-based calibration stage after each of the detection stages in the cascade Merit The proposed method runs at 14 FPS on a single CPU core for VGA-resolution images and 100 FPS using a GPU
Possible Improvements	Method can be used for making image resolution high; here it was used for the face detection.

Research Paper: 008

Title :	"Recognizing Basal Cell Carcinoma on Smartphone-Captured Digital Histopathology Images with Deep Neural Network "
Publication :	Elseviewer, CAMS Innovation Fund for Medical Sciences(CIFMS-2017)
Authors:	Y.Q. Jiang, J.H. Xiong, H.Y. Li, X.H. Yang, W.T. Yu, M. Gao, X. Zhao, Y.P. Ma, W. Zhang Chinese Academy of Medical Sciences, Nanjing, China

LEARNING

Methods Used	The MOI based and WSI based models
Merits/ Demerits/ Outcomes	The MOI- and WSI-based models achieved comparable AUCs around 0.95 Recognizing BCC through smartphone could be considered as a future clinical choice

Possible Improvements	• Proposed method is used for the recognition of skin cancer through smartphone camera image. • Further it can be modified with high resolution camera images • Or else we can use this method for another skin disease.

Research Paper: 009

Title :	"Motion Artifact Reduction Using a Convolutional Neural Network for Dynamic Contrast Enhanced MR Imaging of the Liver"
Publication :	Elseviewer, Japanese Society for Magnetic Resonance in Medicine (2018)
Authors:	DaikiTamada, Marie-LuiseKromrey, Shintaro Ichikawa, Hiroshi Onishi, UtarohMotosugi Department of Radiology, University of Yamanashi, Yamanashi, Japan

LEARNING

Methods Used	multi-channel convolutional neural network based method
Merits/ Demerits/ Outcomes	The trained network was applied to the acquired images to analyze the filtering performance, and the intensities and contrast ratios before and after denoising were compared via Bland–Altman plots Merit improve the quality of images obtained via dynamic contrast enhanced MRI (DCE-MRI)
Possible Improvements	• Proposed method is used for the recognition of Liver related problems • Further it can be modified for another part of the body

	• Or else we can use this method for another criterion like image recognition.

Research Paper: 010

Title :	“Evaluation of fish feeding intensity in aquaculture using a convolutional neural network and machine vision”
Publication :	Elseviewer, Aquaculture (2019)
Authors:	Chao Zhou, DamingXu, Lan Chen, Song Zhang, Chuanheng Sun, Xinting Yang, Yanbo Wang Beijing Research Center for Information Technology in Agriculture, Beijing, China

LEARNING

Methods Used	An automatic method for grading fish feeding intensity based on a convolutional neural network (CNN) and machine vision is proposed to evaluate fish appetite
Merits/ Demerits/ Outcomes	Merit Grading accuracy reached 90%
Possible Improvements	• The training set was still limited due to various conditions • For example, the quantity and consistency of labeled samples. • Designing training algorithms that aremore suitable for small sample datasets

Research Paper: 011

Title :	“Influence of image quality on the identification of psyllids using convolutional neural networks”

Publication :	Elseviewer, Biosystem Engineering (2019)
Authors:	Jayme G.A. Barbedo, Guilherme B. Castro Embrapa Agricultural Informatics, Av. Andre Tosello, Brazil

LEARNING

Methods Used	Squeezenet CNNs and a 10-fold cross validation Strategy
Merits/ Demerits/ Outcomes	A total of 1276 images were used in the experiments, half acquired using a flatbed scanner and half acquired using two different brands of smartphones. Merit Accuracies ranged from less than 70% using only scanned images, to around 90% when only smartphone images were employed
Possible Improvements	• The development of new strategies for image database construction, • For example using the concepts of social networks to involve more people and scale up the efforts towards a representative database • It may be carried out in the context of plant disease identification

Research Paper: 012

Title :	"Vibration-based structural state identification by a 1-dimensional convolutional neural network"

Publication :	Wiley, Computer-Aided Civil and Infrastructure Engineering (2019)
Authors:	Youqi Zhang, YasunoriMiyamori, Shuichi Mikami, Takehiko Saito Department of Civil and Environmental Engineering, Kitami Institute of Technology, Hokkaido, Japan

LEARNING

Methods Used	Simple One DimensionalCNN
Merits/ Demerits/ Outcomes	Merit Very high accuracies 99.79%, 99.36%, and 97.23% in the T-shaped beam, short steel girder bridge, and long steel girder bridge experiments, respectively
Possible Improvements	Method may be proposed for the different structure like wall, railway tracks, etc

Research Paper: 013

Title :	“Windows Malware Detector using ConvolutionalNeural Network based on Visualization Images”
Publication :	IEEE Transactions on Emerging Topics in Computing (2019)
Authors:	Shiva Darshan S.L and Jaidhar C.D Department of Information Technology, National Institute of Technology Karnataka, Surathkal, Mangalore

LEARNING

Methods Used	Simple One Dimensional CNN
Merits/ Demerits/ Outcomes	Detects tiny local structural stiffness and mass changes, and validates the proposed CNN on actual structures The convolutional kernels and outputs of the convolutional and max pooling layers are visualized and discussed as well
Possible Improvements	Method may be proposed for the different structure like wall, railway tracks, etc

Research Paper: 014

Title :	"Reducing the Hausdorff Distance in Medical Image Segmentation with Convolutional Neural Networks"
Publication :	IEEE (2019)
Authors:	DavoodKarimi, and Septimiu E. Salcudean, Fellow, IEEE Department of Electrical and Computer Engineering, University of British Columbia, Vancouver, Canada

LEARNING

Methods Used	Loss functions for training convolutional neural network (CNN)-based segmentation methods with the goal of reducing HD
Merits/ Demerits/ Outcomes	Proposed loss functions can lead to approximately 18- 45% reduction in HD without degrading other segmentation performance criteria such as the Dice similarity coefficient.

Possible Improvements	• To the best of our knowledge, this is the first work to aim at reducing HD in medical image segmentation. • The methods presented may be improved in several ways. • Faster implementation of the HD-based loss functions and more accurate implementation of the loss function based on morphological erosion would be useful. • Moreover, extension of the methods for other applications such as vessel segmentation could also be pursued.

Research Paper: 015

Title :	"Convolutional nets for reconstructing neural circuits from brain images acquired by serial section electron microscopy"
Publication :	Current Opinion in Neurobiology (2019)
Authors:	Kisuk Lee, Nicholas Turner, Thomas Macrina, Jingpeng Wu, Ran Lu, H. Sebastian Seung Department of Brain and Cognitive Sciences, MIT, Cambridge, USA

LEARNING

Methods Used	Neural circuits, Convolutional nets
Merits/ Demerits/ Outcomes	Convolutional nets are also being employed for other tasks in neural circuit reconstruction: Finding synapses and identifying synaptic partners, extending or pruning neuronal reconstructions, and aligning serial section images to create a 3D image stack.

Possible Improvements	• New SEM approaches are being proposed for acquiring such exascale datasets. • Combining SEM and TEM methods with convolutional networkspursued.

Research Paper: 016

Title :	"Medical Image Retrieval Based on Convolutional Neural Network and Supervised Hashing"
Publication :	IEEE Access (2019)
Authors:	YIHENG CAI, YUANYUAN LI, CHANGYAN QIU, JIE MA, AND XURONG GAO School of Information and Communications Engineering, Beijing University of Technology, Beijing, China

LEARNING

Methods Used	content-based medical image retrieval (CBMIR) framework using CNN and hash coding
Merits/ Demerits/ Outcomes	Merit The proposed algorithm combining a Siamese network with the hash method is superior to the classical CNN-based methods. The application of a new loss function can effectively improve retrieval accuracy.
Possible Improvements	The map values of the proposed method are higher than those of other methods, showing that the proposed method is effective in further improving the accuracy of image retrieval

Research Paper: 017

Title :	“A Style Image Confrontation Generation Network Based on Markov Random Field”
Publication :	Springer-Verlag GmbH Germany, part of Springer Nature (2019)
Authors:	GuopengQiu, Jianwen Song and Lilong Chen Strait Animation Institute of Sanming University, Sanming, China

LEARNING

Methods Used	Network based on Markov random field method for image style
Merits/ Demerits/ Outcomes	Merits • Fast style migration, Improve the quality of image • Markov Random Field (MRF) Loss to calculate the image loss, both using the extraction ability of CNN abstract features
Possible Improvements	Improvement in focus on the application in video stream transfer styleMightbe used for other then Van Gogh’s“Star” style.

Figure2.7: Figure from same mentioned research paper[21]

Left figure (content image), the content of the picture in Middle figure is Van Gogh’s “Star” style (style image) as a picture, Right figure is the output in the pictures on left one with middle one as the style of image style after the migration

Research Paper: 018

Title :	“Self-Organizing Maps with Convolutional Layers”
Publication :	Springer Nature Switzerland AG 2020 (2019)
Authors:	Lars Elend and Oliver Kramer Computational Intelligence Group, Carl von Ossietzky University of Oldenburg, Germany

LEARNING

Methods Used	Self-organizing maps (SOMs)
Merits/ Demerits/ Outcomes	Introduce SOM quality measures and analyze the new approach on two benchmark image data sets considering diff erent convolution network levels.
Possible Improvements	• To consider semantic information between labels for evaluation of the SOM results. • To apply the approach to semi-supervised data sets with partially labeled patterns and analyze the learning results

Research Paper: 019

Title :	“Perturbing Convolutional Feature Maps with Histogram of Oriented Gradients for Face Liveness Detection”
Publication :	Springer Nature Switzerland AG 2020 (CISIS 2019/ICEUTE 2019)
Authors:	Yasar Abbas Ur Rehman, Lai-Man Po, Mengyang Liu, ZijieZou, and WeifengOu Department of Electronic Engineering, City University of Hong Kong, Kowloon, Hong Kong

LEARNING

Methods Used	Perturbing the convolutional feature maps with Histogram of Oriented Gradients (HOG) features.
Merits/ Demerits/ Outcomes	Anti-spoofing databases, like OULU-NPU, CASIA-FASD and Replay-Attack, in both intra-database and cross-database scenarios. Merit Experimental results indicate that the proposed framework perform significantly better compare to other approaches in both intra-database and cross-database face anti-spoofing scenarios.
Possible Improvements	Evaluation of perturbing the convolutional layers of CNN using various hand-crafted features, such as Local Binary Patterns (LBP), Shearlet features and wavelet features.

Research Paper: 020

Title :	“Computational neural network in melanocytic lesions diagnosis: artificial intelligence to improve diagnosis in dermatology”
Publication :	European Journal of Dermatology 2019
Authors:	Selim ARACTINGI, Giovanni PELLACANI Department of Dermatology, Paris

LEARNING

Methods Used	convolutional processing combined with artificial intelligence or neural networks (CNN/ANN)
Merits/ Demerits/ Outcomes	Merit Automated diagnostic systems using a CNN improves the management, particularly when using newly developed devices
Possible Improvements	Might be able to achieve performance on par with clinicians across both tasks, demonstrating AI as being capable of classifying skin cancer lesions with a level of competence comparable to that of dermatologists

Research Paper: 021

Title :	"Fully Hyperbolic Convolutional Neural Networks"
Publication :	IEEE 2019
Authors:	Keegan Lensink, Eldad Haber and Bas Peters The University of British Columbia, Vancouver, Canada,Xtract AI, Vancouver, Canada, Computational Geosciences Inc, Vancouver, Canada

LEARNING

Methods Used	reversible CNNs Discrete Wavelet Transform

Merits/ Demerits/ Outcomes	Merit A transform has a natural property that it increases the number of channels while reducing the resolution, while conserving all the information in the image due its inevitability.
Possible Improvements	3D, fully reversible networks will be a key in efficient training and inference.

Research Paper: 022

Title :	"A multi-path 2.5 dimensional convolutional neural network system for segmenting stroke lesions in brain MRI images"
Publication :	Neuroimage Clinical, Springer 2019
Authors:	YunzheXue, Fadi G. Farhat, Olga Boukrina, A. M. Barrett, Jeffrey R. Binder, Usman W. Roshan, William W. Graves

<u>LEARNING</u>

Methods Used	CNN as KF and MCW Images
Merits/ Demerits/ Outcomes	Automatic identification of brain lesions from magnetic resonance imaging (MRI) scans of stroke survivors Training on the KF and MCW images and testing on the ATLAS images yielded a mean Dice coefficient of 0.54. Merit This was reliably better than the next best previous model, UNet, at 0.47.

Possible Improvements	this re-allocation of expert resources will help advance the pace at which we can further our understanding of the critical neural basis of thinking and behavior

Research Paper: 023

Title :	“Adaptive Weighting Depth-variant Deconvolution of Fluorescence Microscopy Images with Convolutional Neural Network”
Publication :	Springer 2019
Authors:	YunzheXue, Fadi G. Farhat, Olga Boukrina, A. M. Barrett, Jeffrey R. Binder, Usman W. Roshan, William W. Graves

LEARNING

Methods Used	Adaptive weighting, convolutional neural network, deconvolution, uorescence microscopy
Merits/ Demerits/ Outcomes	• DelpNet to estimate the defocuslevel of an image patch • The identification accuracy of 98.2% was achieved, which is higher than the accuracy of 95% in based on the same original dataset
Possible Improvements	By using other training sequences we can observe percentage achievement.

Research Paper: 024

Title :	“Motion Artifact Reduction Using a ConvolutionalNeural Network for Dynamic Contrast Enhanced MR Imaging of the Liver”

Publication :	Magnetic Resonance in Medical Sciences, Elsevier 2019
Authors:	DaikiTamada, Marie-LuiseKromrey, Shintaro Ichikawa, Hiroshi Onishi and UtarohMotosugi

LEARNING

Methods Used	Multi-channel convolutional neural network-based method
Merits/ Demerits/ Outcomes	• significantly reducing the magnitude of the artifacts and blurring • A deep learning-based method for removing motion artifacts in images obtained via DCE-MRI of the liver was demonstrated and validated
Possible Improvements	The method can be used for Aqua Images or other then medical images.

Research Paper: 025

Title :	"Underwater Image Enhancement With a DeepResidual Framework"
Publication :	IEEE Access, Advanced Optical Imaging for Extreme Environments 2019
Authors:	PENG LIU, GUOYU WANG, HAO QI, CHUFENG ZHANG, HAIYONG ZHENG, AND ZHIBIN YU

LEARNING

Methods Used	cycle-consistent adversarial networks (CycleGAN) very-deep super-resolution reconstruction model (VDSR)

Merits/ Demerits/ Outcomes	• The loss function and training mode are improved • asynchronous training mode is also proposed to improve the performance of the multi-term loss function
Possible Improvements	Proposed methods to the similar domains, such as image dehazing and super-resolution reconstruction to test the generality of the proposed methods

Research Paper: 026

Title :	“A convolutional neural network regression for quantifyingcyanobacteria using hyperspectral imagery”
Publication :	Elsevier, Remote Sensing of Environment 2019
Authors:	JongCheolPyo, HongtaoDuan, SangsooBaek, Moon Sung Kim, TaegyunJeon, Yong Sung Kwon, Hyuk Lee, Kyung Hwa Cho

<u>LEARNING</u>

Methods Used	Convolutional neural network (CNN) phycocyanin (PC) and chlorophyll-a (Chl-a)
Merits/ Demerits/ Outcomes	CNN regression has the potential to detect and quantify cyanobacteria with high accuracy and can be an alternative to bio-optical algorithms.
Possible Improvements	Therefore, this study will provide the preliminary information for futuredeep learning regression in water quality determination.

Research Paper: 027

Title :	"A Novel Real-time Driver Monitoring System Based on Deep Convolutional Neural Network"
Publication :	IEEE, Instrumentation and Measurement, *2019*
Authors:	Yiheng Zhao, AbdelhamidMammeri, AzzedineBoukerche

LEARNING

Methods Used	refined network HeadNet CNN
Merits/ Demerits/ Outcomes	A high detection accuracy of 96.2% to detect head, and 91.5% to detect facial features, with a fast detection speed of 18 frames per second.
Possible Improvements	• finding a way to weaken the dependency of prediction accuracy towards the imagequality and label measurement • To introduce time based neural network, such as RNN to increase detection accuracy

Research Paper: 028

Title :	"Rotation equivariant and invariant neural networks formicroscopy image analysis"
Publication :	Bioinfomatics,ISMB-ECCB, OXFORD 2019
Authors:	Benjamin Chidester, Tianming Zhou, TianmingZhouMinh N. Do and Jian Ma

LEARNING

Methods Used	CFNet and G-CNN as compared to a standard CNN
Merits/ Demerits/ Outcomes	CFNet has the potential to improve many high-throughput microscopy image analysis applications

Possible Improvements	• finding a way to weaken the dependency of prediction accuracy towards the imagequality and label measurement • To introduce time based neural network, such as RNN to increase detection accuracy

Research Paper: 029

Title :	"Creating the Black Box: A Primer on Convolutional Neural Network Use in Image Interpretation"
Publication :	Elsevier, Current Problems in Diagnostic Radiology2019
Authors:	Toshimasa Clark, Eric Nyberg University of Colorado Anschutz Medical Campus, Aurora, CO

LEARNING

Methods Used	Convolutional Neural Network Use in Image Interpretation
Merits/ Demerits/ Outcomes	This technique is relatively simple to implement, has been shown to demonstrate equivalent performance to neural networks specifically trained on medical image data, and offers a chance for the interested-but-intimidated radiologist to deep her toe in the water without becoming overwhelmed
Possible Improvements	• Finding a way to weaken the dependency of prediction accuracy towards the imagequality and label measurement • To introduce time based neural network, such as RNN to increase detection accuracy

Research Paper: 030

Title :	"Speckle denoisingin optical coherence tomography images using residual deep convolutional neural network"

Publication :	Springer, Multimedia Tools and Applications 2019
Authors:	NehaGour, PriteeKhanna PDPM Indian Institute of Information Technology, Design and Manufacturing, Jabalpur, India

LEARNING

Methods Used	Optical Coherence Tomography (OCT) Evaluated on Duke (SD-OCT) and Topcon (3D-OCT) image databases
Merits/ Demerits/ Outcomes	• It is observed that the proposed approach performs better as compared to the methods referred from literature on both visual and parametric evaluations. • Training of CNN has high computational complexity • An average 3.2 second are required for one OCT image during testing
Possible Improvements	• Average seconds for processing can be less • Training sequence can be modified. • Image training data can be changed and modifies.