World Journal of Advanced Research and Reviews

eISSN: 2581-9615 CODEN (USA): WJARAI

Cross Ref DOI: 10.30574/wjarr

Journal homepage: https://wjarr.com/

(Review Article)

A multi-modal CNN framework for integrating medical imaging for COVID-19 Diagnosis

Mohit Jain * and Adit Shah

University of Illinois, Urbana Champaign, USA.

World Journal of Advanced Research and Reviews, 2020, 08(03), 475–493

Publication History: Received on 15 October 2020; revised on 24 November 2020; accepted on 26 November 2020

Article DOI: https://doi.org/10.30574/wjarr.2020.8.3.0418

Abstract

Due to COVID-19 spreading fast, traditional methods have revealed many inadequacies, showing there is a strong need for better and faster tests. In this article, we discuss the structure, arrangement and clinical use of a multi-modal CNN framework for including medical imaging in COVID-19 diagnosis. When data from X-rays, CT scans and ultrasound are combined, it helps doctors better understand how a disease is showing up in the body. CNN, using each imaging technique's special features, fuses, extracts data and integrates with attention to help diseases be identified more accurately. The article explores how CNN functions, reviews its training approaches and reviews relevant evaluation methods like accuracy, precision, recall, F1-score and AUC. Case studies are shown such as how technology is used in hospitals, for faster triage decisions and remote medical diagnosis. Using case studies and benchmark data sets consistently demonstrates that the model performs better than past, single-technique diagnostic techniques. Problems linked to a lack of data, mislabeling, understanding models and the ethics of AI in healthcare are all discussed. Future work aims to integrate MRI and PET imaging and allow numerous centers to co-operate securely through federated learning. Thanks to the multi-modal CNN framework, medical experts have access to a powerful new tool that not only fights COVID-19 but handles numerous other difficult diseases as well. Merging AI with medical imaging highlights how precision medicine of the future will use machine learning to boost doctor's decisions and help patients faster.

Keywords: Multi-Modal CNN; COVID-19 Diagnosis; Medical Imaging; Deep Learning; AI In Healthcare

1. Introduction

1.1. The Effect of COVID-19 around the World on Healthcare

The outbreak of COVID-19 greatly changed healthcare systems everywhere. Hospitals and clinics often worked to full capacity, lacking beds, devices, and perhaps most importantly, quick ways to diagnose the disease. Quickly diagnosing was vital to helping single patients and slowing down the virus across society.

RT-PCR was the main method doctors used to diagnose COVID-19 at the start. Without exception, routine testing was slow, likely to give false results, and unavailable to people in many parts of the world for unrelated reasons. So, these providers began looking for different screening methods to allow them to make patient triage decisions quickly.

Thanks to medical imaging, the main allies were chest X-rays and CT scans. These techniques let radiologists find signs of COVID-19-induced pneumonia despite RT-PCR test failure or long waits for results affecting medical response. Once there were too many patients for the existing team, the demand for interpreting images outpaced what the team could provide. Managing the many diagnostic cases became easier for clinicians when they started using more technology.

* Corresponding author: Mohit Jain.

The pandemic made it obvious that many traditional systems are especially at risk when so much depends on being done manually. Using diagnostic tools that scaled, gave honest results, and could be counted on in real-time was crucial. Because of the global COVID-19 crisis, mixing artificial intelligence and deep learning into frontline care suddenly became essential.

1.2. The Role of AI and Deep Learning in Pandemic Response

Artificial Intelligence (AI), especially deep learning, became a leading trend during the pandemic. As more and more were required of diagnostics, predictive tools, and tools that scaled, AI made it possible to examine huge amounts of data quickly and accurately. Focusing on this type of network, CNNs have collectively found great success in analyzing medical imaging.

In the first months of the pandemic, scientists and developers used thousands of chest X-rays and CT scans to teach CNNs to find COVID-19 patterns. It was soon seen that these models could match and, in some cases, excel at recognizing patients with COVID-19. CNNs might see tiny signs of lung involvement in disease, which even seasoned radiologists might overlook during stress or when there is lots of work.

AI's benefits reach further than how it reads images. It greatly helped doctors in triage, sorting patients by risk so they don't experience decision fatigue. Because of predictive models, it became possible to estimate when ICUs would be needed and when possible, outbreaks might happen, helping authorities use their resources where needed most.

In COVID-19, technologies based on deep learning shifted from being tested in universities to becoming used in daily operations. Organizations used automated AI systems to pick out COVID-19 cases for clinical teams during treatment. In places with a limited supply of radiologists or very busy healthcare providers, these models provided an urgently needed increase in diagnostic capability.

2. Learning about Multi-Modal Imaging

2.1. What is Multi-Modal Medical Imaging?

Multi-modal medical imaging is a new diagnosis technique that combines different picture methods to give a clear picture of a patient's condition. Multiple exams use different scanning techniques because each combination produces unique and complementary information. Because of this combination, diagnoses become more reliable, we learn more about the condition, and planning better treatments becomes easier.

You can find that every medical imaging modality has advantages and weaknesses. For example, X-rays are often done quickly, can be found nearly anywhere, and are very good at finding issues with bones and lungs. Even so, soft tissue changes that are difficult to spot are hard to find with these techniques. Conversely, the detail and contrast seen with CT scans enable clinicians to study important structures inside the body. While ultrasound gives us real-time pictures, it helps monitor organs and blood by some operators more than others, depending on the part examined.

The important feature of multi-modal imaging is that it combines all the unique advantages of each modality into just one diagnostic system. If physicians combine these images using visual or computational methods, they can better understand the patient's problem and location. It is especially important to show all the disease details when different imaging techniques are necessary.

In these models, multi-modal imaging gives a major boost. Multimodal image analysis with a neural network can allow us to discover relationships not easily found in data from a single source. With these models, texture, density, and shape details from several imaging approaches are understood better, providing stronger and more precise predictions.

Imaging with more than one approach isn't mostly about advancements—it also reflects a change in thinking about how medicine is practiced. It means a healthcare approach that sees the whole picture and notes that dealing with complex diseases usually means paying attention to several factors. As a result of this method, medicine now moves closer to offering precision diagnostics: clear and detailed results that help improve patient outcomes.

2.2. Common Imaging Modalities Used in COVID-19 Diagnosis

In the COVID-19 pandemic, doctors looked to medical imaging for fast diagnosis, checking disease progression, and rating the degree of lung involvement. People have turned to three common imaging methods: chest X-rays, CT scans, and lung ultrasounds. These modalities brought something important that made them useful at various times and locations throughout the COVID-19 experience.

X-rays of the chest were one of the best and most broadly applied resources during the pandemic. Since they're portable, doctors could use them in treating patients in critical care or during isolation. Among the typical findings of COVID-19 on X-rays are bilateral infiltrates, ground-glass opacities, and consolidation. Nevertheless, given mild or early cases, their sensitivity was weaker than that of CT scans. However, because emergencies required fast treatment, these drugs had to be used anyway.

CT scans gave us much better sensitivity and detailed information about the body's anatomy. Fine details in the lung tissue could be captured clearly on high-resolution CT images when looking at COVID-19 pneumonia. Typical CT findings among COVID-19 patients were ground-glass opacities, enlarged vessels, and thicker interlobular spaces, frequently close to the lungs' outer regions or the back of the lungs. By looking at the clear images, clinicians sorted patients by how much their lungs were damaged, changed their breathing management, and monitored the effects of treatment.

Before the pandemic, ultrasound was seldom applied to imaging the lungs; its rise now depends on its portability, safety without radiation, and capacity to do tests at the patient's side. Using POCUS at the bedside helped us immediately view pleural and lung-related issues, proving most beneficial in the ICU for watching changes in fluid. Using a lung ultrasound, you can find B-lines, thickening of the pleura, and subpleural consolidations linked to COVID-19. Growing importance was placed on this method to help when imaging needed to be fast, repeatable, and safe.

Every imaging technique was important in the effort to control the COVID-19 pandemic. CT scans resolved every detail to assess the lungs, X-rays sped up initial triage, and ultrasound allowed for quick bedside monitoring. When used together in a deep learning framework, such different modalities could detect disease much more reliably.

When CT, X-ray, and ultrasound images are used in a single deep-learning system, the results help clinicians achieve more confidence in their diagnoses. Because CT images show structure, X-rays give immediate results, and ultrasound is dynamic, physicians can view all disease sides. This approach improves opportunities for early detection, fewer mistaken negative test results and clinical decisions become easier. With COVID-19, multi-modal imaging isn't mainly about better user experience—it has become necessary for patient care.

Table 1 Comparison of Imaging Modalities for COVID-19 Diagnosis

Imaging Modality	Strengths	Weaknesses	Common Findings in COVID-19
Chest X-ray	Fast, widely available	Low sensitivity, overlapping structures	Bilateral infiltrates, patchy opacities
CT scan	High resolution, detailed lung anatomy	Expensive, high radiation, limited access	Ground-glass opacities, consolidation
Ultrasound	Portable, real-time imaging	Operator-dependent, limited depth	B-lines, pleural thickening

3. Convolutional Neural Networks (CNNs)

3.1. Basics of CNN Architecture

CNNs, convolutional neural networks, are carefully built to handle and inspect pictures and videos. Unlike other neural networks, CNNs are great at noticing patterns in images, so they have many uses for medical imaging, like locating, discovering, and classifying diseases.

At the center of any CNN is a convolutional layer where filters are used on the input image. Running a convolution over the filters on the image, they can pick up levels of detail and identify edges, corners, and textures. During the many stages of the network, more complex patterns, such as shapes and structures, appear and become crucial for examining medical images.

Basic elements in a CNN architecture are

- Convolutional Layers: These layers use several filters to find the main features contained in the input image. Every filter gathers a unique image, helping the network get a full picture of its content.
- Activation Functions: ReLU is normally used after each convolution to make the result more complex. With this, the model learns difficult patterns by copying how neurons become active in the brains of humans.
- Pooling Layers: These layers lessen the size of the data's spatial dimensions by gathering functional details from them. Max pooling remains the key feature (the maximum value) from each region, which supports faster computation and prevents overfitting.
- Fully Connected Layers: This layer turns the extracted information into predictions. The neural network can make its final call because every neuron in a fully connected layer is connected to all those in the previous layer.
- Dropout layers: are used only during training since they make some neurons active or inactive to improve how the model can generalize new information.
- Output Layer: Typically uses a softmax or sigmoid function, depending on whether the problem is about many classes or only a few classes.

Automatic feature extraction is one of the strongest abilities CNNs possess. While existing image analysis methods depend on useful features made with specific knowledge, CNNs can identify those useful features from the data during their training. That's why they often work better in healthcare, because the smallest details in medical images may guide treatment options.

Thanks to their structure, CNNs are very effective for image analysis, and they're often found in medical images such as chest X-rays, CT scans, and MRIs. To work well, they must be trained with lots of tagged data, but when well-trained, they can quickly sort through and analyze any image with high accuracy.

Overall, CNNs combine mathematical brilliance with inspiration from nature to help machines read images and draw growing levels of abstraction from them. Using this feature makes AI-based imaging solutions effective, particularly against COVID-19.

Table 2 CNN Architectures Used in COVID-19 Imaging Studies

Architecture	Layers	Modality Used	Performance (Accuracy)
ResNet-50	50	Chest X-ray	93%
DenseNet-121	121	CT scan	91%
Inception-V3	48	Multi-modal	95%

3.2. What makes CNNs well-suited for medical imaging

Convolutional Neural Networks are the leading method in medical image analysis for several reasons. Their special abilities and building structure are exactly what medical scans need when interpreting scans used to diagnose COVID-19. Yet, what gives CNNs their unique advantage for medical imaging?

In the first place, medical imaging commonly focuses on examining detailed and very clear photographs. Some details or irregularities within the images can be very important for diagnosis. CNNs are great at spotting and learning from the tiny details within images. A naked-eye review of the results would likely miss the sizes and shapes of problems in lungs seen with AI tools.

In addition, CNNs achieve useful results by recognizing spatial hierarchies. The structure of most medical images reflects that an organ or tissue has its shape, texture, and position around or within a patient's body. Throughout the network, CNNs preserve the connections between different parts of any image. As a result, it becomes easier to identify where the problems are within the patient.

CNNs also stand out because they are tough to change in image quality. A medical scan can change a lot depending on what equipment is available, the person's physique, or the temperature in the room. Using different datasets, CNNs can learn to handle these inconsistencies and maintain reliable diagnostic results in all types of settings and with all populations.

Because CNNs have many layers, they are easy to scale. They can be taught to work with 2D images, such as X-rays and 3D images from computerized tomography (CT) or magnetic resonance imaging (MRI). They work well in several kinds of imaging and various patient uses.

CNNs allow diagnostics to be done automatically. Many more patients survived thanks to medications and care from hospitals, but too many cases for radiologists and hospitals were more than they could handle, especially cases of COVID-19. The use of CNNs enabled the initial scan to be automatic, making diagnosis happen more rapidly and giving preference to patients at greatest risk.

Unlike other networks, CNNs can also use data from many different sources. It is possible to design CNNs to harvest information from various imaging sources (CT and X-ray) and include things like a patient's history and lab results. As a result, diagnostic accuracy is boosted due to the use of a wider diagnostic framework.

CNNs are ideal for carrying out analyses that happen in real-time. Following training, a CNN can quickly analyze newly received images, so it is excellent for making quick judgments in emergency and intensive care environments. During the COVID-19 surge, having such an ability was very important for making prompt discoveries.

4. The Need for Multi-Modal Integration in COVID-19 Diagnosis

4.1. Problems with Depending on a Single Type of Analysis

Although only using X-rays or CT scans can be valuable in treating COVID-19, some major problems can limit their effectiveness. The significance of these limitations is most clear when you use just one scan type since COVID-19 can affect various body parts and has different signs and symptoms.

X-rays are easy to obtain, inexpensive, and popular, but they cannot find the earliest or mildest signs of lung infection. Because milder forms of COVID-19 are not always obvious on an X-ray, the test may not catch the virus in time for the proper treatment. Moreover, it's common for organized structures and reduced differences to hide important parts, something that often occurs with portable or bedside imaging in intensive care units.

According to research, COVID-19 pneumonia is often more clearly visible on CT images than on other scans. Even so, CT imaging does have some problems. The tools are pricey and usually hard to find in rural or low-resource areas, so moving patients to these tools and giving them radiation is needed. CT scans sometimes report abnormalities that seem like COVID-19 but later result from other viral or bacterial infections.

The usefulness of ultrasound in dynamic monitoring is strong, although it relies much on the experience of the person using it. Differences in using and reading images can cause inconsistent findings, mostly among users with little experience. Infections that spread to far parts of the lung can be missed by ultrasound since ultrasound images do not penetrate deeply enough.

Using an imaging modality alone puts staff at risk of missing or mistaking a diagnosis. Using only one technique will not give you the whole story. An individual's age, previous diseases of the lungs, and how far along the infection is can cause differences in imaging results. For this reason, depending on a single imaging method may not provide enough accurate information.

COVID-19 made it obvious how much healthcare is lacking in certain parts of Europe. High numbers of infections in outbreaks made it obvious that accurate, quick, and scalable diagnostics were out of reach for just one type of diagnostic imaging. Because COVID-19 symptoms range from nothing to severe breathing issues, it was necessary to deal with patients using several strategies at once.

As a result, a diagnostic gap remains, which can only be resolved by using more imaging techniques. Having extra images is useful, but a multi-modal technique adds explanation, depth, and trust to the diagnoses made by radiologists. That broad approach is vital for achieving accurate, fast, and life-saving answers during the pandemic.

4.2. Reasons for Using Many Imaging Sources

Blending multiple multimodal imaging tools dramatically improves accuracy in diagnosing complicated and swiftly developing diseases like COVID-19. Because X-ray, CT scan, and ultrasound data are combined, doctors and artificial intelligence tools can make better and quicker diagnoses.

Multi-modal integration brings about one of its most important benefits: complementarity. All imaging modalities give doctors a different view of what's happening inside the body. Chest X-rays are great for early lung involvement detection, but a CT scan makes it possible to find where and how severe ground-glass opacities or consolidations are. Unlike X-rays, ultrasound makes it easy to see how the pleura moves and how much fluid is present at any time. When these strengths are joined, multi-modal systems guarantee that all key information is picked up.

Imaging with several methods can improve the certainty of a diagnosis. In many clinical cases, radiologists must decide based on limited or vague information. When images are integrated, information from different kinds of scans can be matched to clear up inconsistencies. A dead giveaway of a suspicious shadow on a chest X-ray can be resolved by CT scanning, which helps to cut down on misdiagnoses.

The use of CNNs, in particular, greatly increases the potential of combining these technologies. CNNs designed to handle multiple types of images can quickly detect complicated designs and overlapping aspects between images that could go beyond what human experts notice. Since neural networks can adjust to context, every input type is used appropriately to create the best and most detailed prediction.

Furthermore, these systems are known for being solid and flexible. How a disease is shown and the patient's response to imaging tests can differ. A CT scan might not happen for a critically ill patient because it could endanger them in transit. We can still get detailed diagnostic information by doing bedside ultrasounds and portable X-rays. Flexibility of use is possible with multi-modal systems, regardless of the radiological tools used in a clinic.

In practice, combining different kinds of health data can enhance how fast and effectively triage is done and how treatments are planned. When clinicians know what is happening with a patient at every level, they can quickly determine who should receive hospitalization, intensive care, or mechanically assisted breathing. It allows for constant disease monitoring, assessment of treatment progress, and better decision-making about using resources.

When a pandemic occurs and fast, large-scale work is needed, multi-modal imaging with AI tools has shown it can help with diagnostics, handle more patients, and ensure reliable, high results. It is a sign of what lies ahead in diagnostics—easing the detection not just of COVID-19 but of many other conditions that need to be seen from several angles.

5. Designing the Multi-Modal CNN Framework

5.1. Data Acquisition and Preprocessing

Data collection and preliminary data treatment forms the foundation of an effective CNN framework and usually decides if the model will be successful. It is very important to get reliable, varied and correctly labeled data when using medical images for the diagnosis of COVID-19 which can be challenging. How correct, useful for patients and useful for wider medical use a model is, depends a lot on the quality and range of the data it processes.

To get the data, samples of imaging such as chest X-rays, CT scans and ultrasound, are taken from COVID-19-positive and COVID-19-negative patients. They must come from different healthcare institutions and places to include a range of kinds of equipment, patients and diseases under study. If patient age, gender, past health conditions and stage of the disease are metadatas in the data, it enriches the model's ability to learn.

All images used in multi-modal training should be both in the same time frame and fit the same clinical phase. The images (CT, X-ray or ultrasound) used for a certain patient should be taken not too far apart and be labeled using the same diagnostic conclusion throughout the process. A model's learning process can be seriously hindered and its rule-making go astray if the data is not coherent.

At the start, the input data is handled using various steps to make it standard and clean. Images are usually modified to the fixed input dimensions; this makes all images the same size for the network. The data is made consistent by scaling the pixel values which helps lessen the difference caused by using different imaging conditions. It improves how well and smoothly the training process works.

Contrast enhancement and denoising filters are normally used to improve the quality of the dataset. They can help spot important features such as opacities, consolidations or thickening of the pleura which might signal COVID-19 on a lung scan. Experts mark areas or examples that may help identify coronavirus such as those seen in images of COVID-19 cases. Because supervised learning needs clear examples, labeled inputs make the model focus on parts of the data that are important in medicine.

Data augmentation is very important for overcoming problems caused by small datasets. By combining rotation, zoom, translation and brightness changes, the training data can be made larger which helps the model learn more generally and reduces the risk it overfits. For the validity of the data and proper data relationships, these augmentations have to be repeated on every data modality when processing the same patient record.

Another factor is how to separate the dataset into parts for training, validation and testing. Ensuring no pictures of the same patient appear in more than one group makes sure there is no leak of information and that performance evaluations are right. It is very important to check the class balance regularly, as COVID-19 data tends to include more cases of illness than those that stay healthy. Many times, using stratified sampling or sampling the minority class multiple times is used to deal with this issue.

5.2. Model Architecture for Fusion

Immediately after preparing the data, the following important step is to design an architecture that properly combines images from different modalities. The point is to put together a network that can analyze the information from various scans and put it together into one integrated interpretation.

When a network involves chest X-rays, CT scans and ultrasounds, each modality is usually handled by its own branch within the network. Every branch counts on convolutional layers to find different features independently on their own input data. Features that can be found include edges, different textures or other patterns that show if there is disease and how severe it might be. The early layers today are designed to recognize features unique to each imaging tool which protects its special abilities in diagnostics.

It is important to know when and in which ways to bring together the different modality-specific aspects. Among all the methods, mid-level fusion proves to be especially successful. Each picture, sound and video is handled separately by several convolution and pooling layers and the resulting feature maps are fused together at a further stage. Every branch is able to keep its own field of study, but also can be understood by all when put together. After fusion, the integrated feature map is used by further convolutional layers to identify and understand the combined details for making one final diagnosis.

Some people are also using attention mechanisms in how their architectures are built. They make it possible for the network to determine which aspects of the information are most important for a diagnosis. If the clarity of an X-ray is less good, the model can put more emphasis on information from CT pictures. Attention mechanisms make the network respond to and cope with new situations, making it stronger in practice.

At the last part of architect architecture, all the features are sent through fully connected layers which bring everything together and deliver a probability for whether COVID-19 is there or not. For special uses, the output from classification could comprise several groups (e.g., mild, moderate and severe infections) or go farther to estimate how the disease might progress or react to treatment.

Dropout layers and batch normalization should be part of the model because they stop overfitting and maintain good learning performance. Because of these parts, the model can handle different cases without requiring retraining which helps in various medical environments.

5.3. Feature Extraction and Integration Strategies

The main steps in a multi-modal CNN are feature extraction and integration which help it change the raw picture data into a reliable clinical report. At this point, it is checked whether the model can grasp, incorporate and make good use of what each imaging test offers.

In the early stages of each branch, the convolutional layers are responsible for the first phase which is feature extraction. They are able to find different types of patterns in the images, checking first for simple edges, then more complicated things such as lungs or disease spots. Health issues such as bilateral opacities or consolidation could be something the algorithm learns to look for in a chest X-ray. Ground-glass opacities, thickened blood vessels and nodules are sometimes found in CT scans with excellent accuracy. Problems such as irregular pleural lines and B-line artifacts might be found by ultrasound technology.

Each methodology gives a separate understanding of the problem. An X-ray shows general lung opacity, a CT scan gives specific information about lung segments and an ultrasound shows live changes in fluid collection. Making sure these distinctive aspects stay intact is very important to keep the growth of the company strong.

After getting the features, the main task is making those features part of one accurate predicting vector. Concatenation is widely used which flattens out each modality's feature maps and stacks them together as a single vector. Applying this approach means no important features are dropped which gives the dense layers space to identify how they all relate in the final stage.

More skilled approaches include working on elements such as summing or multiplying which highlight similar or related elements in all the input data. A new kind of fusion method, called attention, allows the model to decide which features are important for each task. If particular features are longer-lasting in one modality, attention processes make those more prominent to help the degree of conscious thought.

Sometimes, architectures arrange feature extraction and integration to take place at multiple points instead of only only one fusion layer. Thanks to multiple levels of fusion, the model gets better at telling caring for similar health situations apart.

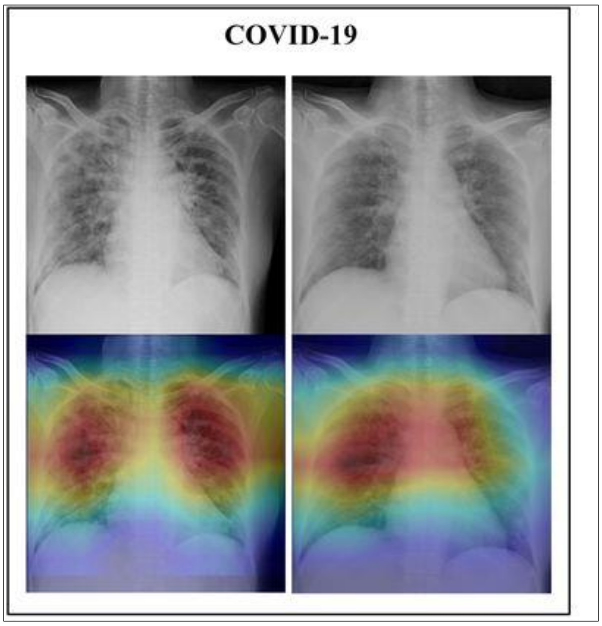

Figure 1 Feature Maps from CNN Layers (Heatmap Visualization)

6. Training and Optimization of the CNN Framework

6.1. Data Augmentation Techniques

Diagnosing COVID-19 using a CNN and multi-modal technology mostly depends on having a wide variety of quality data. It is often troublesome to create big, labeled datasets in this field because of privacy, the complexity of labels, and the lack of interested patients. That's when data augmentation is used—it enlarges the data to improve the model's ability to work in various situations.

Data augmentation describes how we create new images from original images by applying various changes. An attempt is being made to help the model learn under multiple circumstances and minimize its use of particular image elements. Large differences in COVID-19 presentation mean it is vital to be familiar with the disease in imaging.

When a system has multiple output channels, augmentation should be implemented in the same manner throughout to secure their relationship. As a result, when paired CT and X-ray images are transformed, each transformation should be done to the image of the same patient. The study can lead to confused learning if different augmentations are applied simultaneously to several modalities.

Standard processes include rotation to show how a photo may look from other patient angles, flipping horizontally and vertically to help the model recognize any changes in a patient's direction, and zooming, meaning the image might be made bigger or smaller. When you adjust the brightness and contrast, you can fake the variations that result from different ways of capturing or lighting an image. Indeed, in more advanced cases, distorting the model and introducing random disturbances helps it become more robust.

For data on COVID-19, using augmentation plays a key part in ensuring equal numbers of each class. You'll find that most available datasets have more images depicting sickness other than COVID-19 than actual COVID-19 cases. Using oversampling of the minority group with artificial photos, the model recognizes such cases equally and without favoring the majority class.

Generative augmentation is a different approach; networks called GANs (Generative Adversarial Networks) create realistic but fake images. This method is very useful in modeling parts of COVID-19 that do not occur often in actual patient records.

All things considered; data augmentation is required by deep learning algorithms in medical imaging. Data augmentation works like a data enhancer, bringing more variation to the data, stopping the model from overfitting, and making it perform much better when encountering new examples. For this type of multi-modal system, a data set that has been wisely augmented greatly contributes to its effectiveness.

6.2. Hyperparameter Tuning and Loss Functions

As soon as the data is prepped and augmentations are set up, the next step is to fine-tune hyperparameters and choose the right loss function. These two parts of the model greatly determine its learning speed, how quickly it converges, and how accurate its diagnosis becomes.

The settings for a CNN's hyperparameters must be made before data starts moving through it. The example parameters used are the learning rate, batch size, number of epochs, the type of optimizer, how much dropout is used, and how many convolutional filters there are in each layer. Unlike parameters automatically found during training, hyperparameters must be set by us or by another algorithm to make sure the model learns correctly.

The learning rate is the most important of all. It decides how large the model will be and adjusts its parameters when it sees an error. Too high a learning rate might bring the weights past the optimal ones, and too low one will make training barely move forward or stop near suboptimal weights. Training programs often opt for adaptations like those the Adam and RMSprop algorithms offer.

The size of the batch you train is very important as well. Small batch sizes lead to a stochastic gradient in updates, which improves your ability to avoid local minima in the cost of producing training noise. Running more training examples at once stabilizes learning, yet it can use much memory and yield less than ideal solutions. The choice of model depends solely on the number of networks and the dataset size.

It is important to deal with overfitting in multi-modal systems, so dropout and regularization are used to prevent it. As neurons are randomly shut off during training, the model works on representations that can be used in similar situations.

Although you can tune hyperparameters by hand, people normally rely on grid search, random search, or Bayesian optimization. With these approaches, researchers explore all hyperparameter combinations to locate the combination that gives the best results.

Picking the best loss function goes together with selecting hyperparameters to guide the learning in a model. Most often, categorical cross-entropy is chosen for multi-class classes in COVID-19 detection and binary cross-entropy for positive or negative options. The main job of these functions is to compare the predicted outcome with the actual one, telling the network to let backpropagation reduce this difference.

To solve class imbalances, some loss functions are modified to emphasize mistakes involving minority classes. This way, the model won't always guess the dominant class, which can cause much bigger harm in healthcare when patients are not diagnosed.

In multi-modal learning cases, many architectures add functions like segmentation masks or attention scores as extra contributions to control training and boost intermediate results.

7. Evaluation Metrics and Performance Analysis

7.1. Accuracy, Precision, Recall, F1-Score

Evaluating the results using reliable and standard measurements is necessary once the multi-modal CNN is used for COVID-19 diagnosis. These figures offer a view of the model's effectiveness with new data and help explain its strengths and weaknesses to clinicians and those who built it. During a pandemic, the goal of medical diagnosis isn't limited to the accurate model. Instead, the model must handle the complexities of deciding between life and death.

Accuracy is the simplest way to check a model by comparing how many times the model was correct to how many total cases there are. Even though it makes sense, accuracy doesn't always give true results when sets are unequal. If 90% of the images in a dataset are negative for COVID, a model that only predicts negative won't be useful for diagnosis.

At this point, precision starts counting positive results correctly to know how much of the model's predictions are right. We learn from it how many positive cases turned out to be true. Since COVID-19 causes stress, ensuring the model doesn't report false positive cases helps avoid extra worry, tests, or treatments.

On the other hand, recall means that true positive predictions are divided by actual positive cases. It helps to answer the vital question: How many did the model get right among those who do have COVID-19? Because false negatives in testing during a pandemic might mean individuals can keep spreading the virus or lose out on important treatment, having a good recall is particularly important.

We fix this discrepancy in the F1-score, taking the harmonic mean of precision and recall. This metric stands out when the data is uneven, allowing the model to show accuracy for both categories of cases. It is important for any medical tool in healthcare that the model is precise and sensitive about identifying COVID-19, so a high F1 score is crucial.

Many times, metrics are interpreted by each modality and the whole network. Using this approach, researchers find out if the results are improved by using various imaging techniques together and if the model uses the strengths of each image type.

Often, confusion matrices show how a model is doing for every class it classifies data into. You can see the true and false values for both positives and negatives, breaking down how the model behaved during its predictions. Such graphics allow for spotting discriminatory factors in the model and may signal problems in the data.

Table 3 Evaluation Metrics across Models

Model Type	Accuracy	Precision	Recall	F1-Score
Single-Modality (X-ray)	87%	85%	88%	86.5%
Single-Modality (CT)	89%	88%	90%	89%
Multi-Modal CNN	94%	93%	95%	94%

7.2. ROC Curve and AUC Evaluation

Other advanced methods like the ROC curve and AUC allow us to analyze models better, which is useful when evaluating whether a sample is COVID-positive or COVID-negative. These tools show how sensitive and specific the model is to decision threshold changes, allowing us to assess its performance in determining classes differently.

The ROC curve compares the recall ratio and the rate of false positives for all different classification methods. Every step along the ROC curve demonstrates that increasing how many correct discoveries the model can make might raise the false alarm rate. A model that plots its curve in the top left portion of the graph suggests high recall and very few false positives.

Medical imaging is especially helpful for ROC curves since clinical concerns might lead to a change in diagnostic threshold. In a pandemic, recall should take more value than specificity because missing any infected patient is undesirable. With the ROC curve, clinicians and data scientists can determine what combination of sensitivity and specificity makes the most sense in everyday practice.

The AUC shows the total performance of the model through the ROC curve. One number represents the model's power to tell classes apart. If an AUC score is 1.0, the data identifies each group with complete accuracy, and if the score is 0.5, it belongs to the same group as a random guess. In many situations, having an AUC above 0.9 is exceptional; having one between 0.8 and 0.9 is fine, and one between 0.7 and 0.8 works for most uses.

Multimodal CNN models allow the calculation of AUC for each data type and the integrated system. Comparing the different methods will enable us to confirm that using both CT and X-ray helps improve distinguishing performance instead of adding unnecessary information.

With ROC and AUC, model performance can be improved by modifying when to call a positive result. A hospital might ease up on the decision rule to avoid missing infections during heavy pressure or make it more strict when the community's infection rate is low. They allow stakeholders to use AI models according to their values and processes.

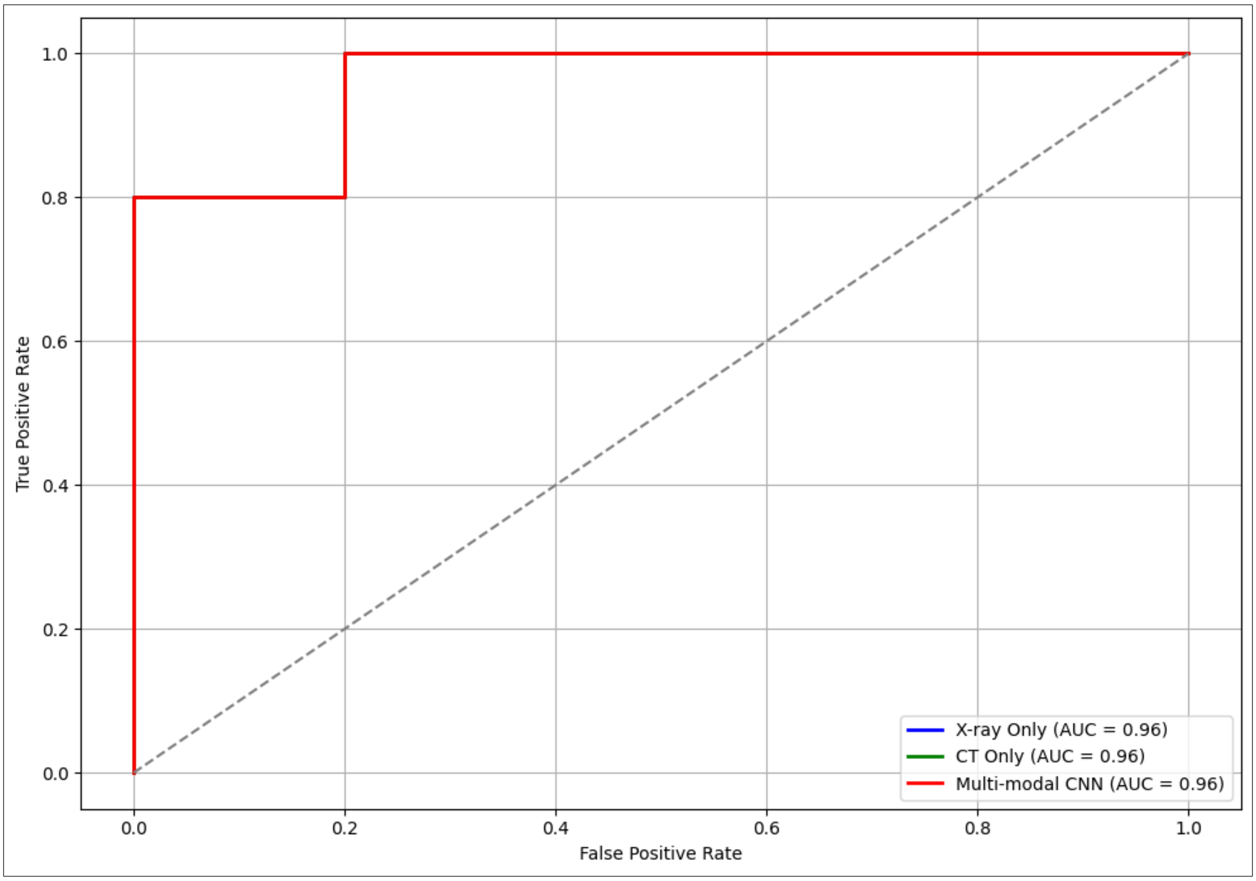

Figure 2 ROC Curve Comparison of CNN Models

8. Case Studies and the results from experiments

8.1. Benchmark Datasets Used

Proper evaluation and practical use of a multi-modal CNN to detect COVID-19 depend greatly on the availability of benchmark datasets. Data in these groups are used for all three types of evaluation, and the model's effectiveness and strength depend on how good and complete the data are. Finding labeled datasets with multiple forms of imaging per patient is a key challenge for anyone working with multi-modal systems.

The COVID-19 dataset became popular during the pandemic by combining chest X-ray images from Cohen's COVID-19 Image Data Collection and the RSNA Pneumonia Detection Challenge. Though it includes numerous X-ray images, COVID-19 is missing matching CT or ultrasound data, so its use in true multi-modal training depends on receiving this extra data from somewhere else.

The COVID-CT dataset gives labeled images from people with COVID-19 and normal controls, all taken from different publications and establishments. Annotations in the dataset explain ground-glass opacities and important signs of COVID-19 pneumonia so that they can be used for supervised learning and segmentation cases.

Applicants can also turn to the SIRM dataset, released by the Italian Society of Medical and Interventional Radiology, which combines X-rays and CT scans. Although there are few standard annotations, many European researchers used this data to train hybrid CNN models.

Certain study teams add details about a patient's age, symptoms, blood oxygen levels, and other health complaints to their datasets. Using metadata in CNNs allows for analysis that goes above and beyond looking at images.

Researchers have included datasets collected by Mount Sinai Health System, Stanford, and several hospitals in Wuhan, China. Since the data is kept private for patient privacy, case studies have been able to use them, often showing how CNN models function in real-world use.

However, although lots of data were released publicly during the pandemic, only a little of it includes paired imaging from different scanners and modalities. Many researchers help address this by unifying different X-ray and CT datasets that share similar disease severity and symptoms or happen simultaneously. Though imperfect, this approach makes learning from data mixes possible since separate datasets are unavailable.

A good or bad general performance from the model depends largely on what kind of data is used. Having a variety of data helps CNN identify differences caused by scanner types, people's backgrounds and healthcare settings, and specific signs of disease. True reliability in a multi-modal model requires benchmark datasets with good diversity, representation, and quality annotation.

8.2. Benchmark Datasets Used

It is necessary to match the performance of the new multi-modal framework with that of older, single-modality, and common methods used in COVID-19 diagnosis. The analysis of results allows us to confirm both the usefulness of multi-modal integration and the readiness of the new system for practical use.

Johnson said that the most trusted diagnostic method for confirmation is still RT-PCR testing for COVID-19. Issues—such as slow reporting, incorrect results from testing not all people, and problems accessing them in locations without resources—have made rapid and trustworthy solutions desirable. Alternatively, medical imaging can be used, though often only scanning with a single mode means we miss some important details.

Alternatively, using a single modality, for example, diffusing deep learning models with ResNet, VGGNet, and DenseNet, accuracy levels of 85% to 95% are often observed on chest X-rays. Yet, these methods frequently have difficulty with cases that are hard to define and fuzzy scans. CT-based CNN models make the findings' details clearer, and they slightly outperform plain film-based models, mostly in sensitivity. However, the models often only work for a limited group of patients.

Many experiments using multi-modal CNN approaches report improved critical evaluation metrics results. A CT-Xray combination study showed that using two methods resulted in an F1-score and AUC improvement of 7–10% compared to the model using only one type of imaging. This increase is possible because by using several learning styles, each one deals with the problems of other forms.

Multi-modal networks generalized well on data from different hospitals when tested using comparative trials. With external tests, models using one data type often performed worse than expected. However, multi-mode frameworks maintained their accuracy and lower overfitting rate thanks to their better range of features.

A further comparison point is how fast and easy the diagnostics are for each tracker. When ready for use, the model gives a diagnosis in just a few seconds and—with heatmaps and attention graphics—helps doctors better understand what led to the prediction. As a result, these techniques give us more confidence and enable the use of these models in real clinical applications, where understandability and trust matter.

Multi-modal frameworks perform better than previous machine learning models that need manually designed features. Classic supervised methods cannot handle large medical image data as well as adaptive deep CNNs can, regardless of how carefully they are tuned.

9. Real-world applications and Use Cases

9.1. Clinical Implementation of Multi-Modal CNNs

Whether an AI system performs well in clinical environments is the main test for any healthcare solution. Effectively developed and validated multi-modal CNN systems may be valuable in real clinical situations, mainly for managing and diagnosing COVID-19. Improvements in computing, useful software, and medical equipment are helping hospitals and centers introduce teleradiology.

The CNN is used on in-house confidential and cloud servers connecting with the hospital's existing imaging systems to make clinical implementation possible. As a result, all X-rays, CT scans, and ultrasound images can be seen instantly. Upon uploading, the CNN takes care of the images and produces a diagnostic report or score in seconds.

Helping with triage is one of the greatest advantages of these systems. When emergencies happen or during a pandemic surge, every second counts. Clinicians can make fast decisions about COVID-19 cases because multi-modal CNNs help diagnose them and decide the next steps. Such systems are most useful in regions where radiologists might not be enough and are occasionally absent.

Generating accurate second-opinion cases is an additional use case for AIs. Eventually, making a decision is important, especially when it's not always clear what the process requires. CNN scores and heatmaps give analysts the data perspective to verify human judgments.

Several hospitals have installed multi-modal CNNs in mobile radiology vehicles or telemedicine systems, allowing remote care and spacing out hospital crowds. This situation benefits rural or underserved areas where experts are hard to find.

In addition, such models are now widely applied to observe how diseases progress. When scans are studied at various times, CNNs can see if COVID-19-induced pneumonia improves or worsens. Because of this, clinicians may choose to use mechanical ventilation or to approve discharge as appropriate.

Hospitals rely on AI models to review past images in their databases, detect possible overlooked diseases, and make their protocols safer for everyone involved.

For those at the cutting edge of healthcare, using multi-modal CNNs is already a reality. Pandemic management is only the start for these tools; they plan to transform support for diagnosing many diseases and types of imaging.

9.2. Integration with Hospital Management Systems

Successful deployment of multi-modal CNN frameworks in hospitals depends on easy integration with HMS. Any system for digital health must link up with other parts of the digital health system to ensure the easy transfer of information and privacy is maintained.

First, it is necessary to link the CNN system to PACS so that radiology images are accessible there. The model has been set up to automatically get chest X-rays, CT scans, and ultrasound images from PACS, check them immediately, and send the results back immediately. Now, the diagnostic process runs faster because there is no need for manual work.

The Electronic Health Record (EHR) system is the next system that Integrated Delivery Networks need to integrate with. The results can automatically appear in the patient's EHR as soon as the model predicts. Having the data next to lab reports, past medical information, and vital signs gives clinicians a better view of what is happening with the patient.

It also allows automated alerting and workflows to start. So, if CNN foresees a high chance of disease, HMS will send a warning to infection control teams, schedule the test, or pass the case to a more experienced radiologist for review. With these systems, important cases are dealt with promptly.

Because HIPAA and similar regulations require it, secure encryption and an audit path must be programmed into the integration. As a result, patient data can be handled quickly, securely, and ethically.

Also, when integrated with HMS, planners can easily review model results, find cases suitable for quality reviews, and extract insights about whole populations. Health information can guide resource usage, make government policies, and advance research.

10. Challenges and Limitations.

10.1. Data Scarcity and Labeling Issues

Multi-modal CNNs can be useful in COVID-19 diagnosis; even so, achieving reliable and robust AI models is very difficult because of the problems with data scarcity and the accuracy of labeling because deep learning values data, gathering and correctly labeling medical images of multiple types is hard.

A major difficulty is that multi-modal data is not widely available in a useful pooled format. Although resources like COVID-19 and COVID-19 comprise one form of data per patient, not many datasets have the same data recorded in multiple ways. Because of the restriction, the model misses important relationships among the different input types, which lowers its effectiveness and usefulness.

Company records can benefit from images, but creating understandable labels for images is still difficult. It's common for expert radiologists to spend lots of time and make subjective judgments by labeling medical photos. This can get even more complicated since clinicians with experience often disagree on varied or ambiguous conditions.

Since COVID-19 keeps changing, it's even harder to apply proper labels. The infection develops in people differently, and the patterns seen on X-rays differ according to the stage of the disease. Because labels should be stamped with time and recognize the circumstances, producing datasets is more complex.

In most cases, there are far greater amounts of normal data than data showing disease. This problem makes the model more likely to treat patients as healthy only, even if some show COVID-19 signs.

Scientists regularly depend on data augmentation and transfer learning to resolve these issues, but those techniques only provide temporary solutions. Lasting answers come from joining forces and ranging hospital and country data, ensuring the data is diverse and large enough. Yet, this means dealing with extra problems about who controls data, how it is used, and how it connects with other systems.

All things considered, CNN's achievements rely critically on how much high-quality data is accessible. Before these data problems are fixed, the full potential for multi-modal diagnosis in COVID-19 and other areas cannot be achieved.

10.2. Model Interpretability and Ethical Concerns

With more complex CNN models coming into medical care, we face a major problem: how they draw their conclusions is becoming less clear. Since these models are complex, the methods used to make decisions are rarely clear. Because healthcare influences people's lives, lacking transparency could lead to big problems.

They do not widely use models that clinicians cannot verify or explain. If a model shows a high risk of COVID-19, doctors shouldn't take it at face value. Is the contrast pattern only found in particular aspects of the CT image? There's something different about the pattern on the X-ray. Applied broadly, these findings could even be used together. The most accurate models may not convince doctors and other health professionals if explanations are unclear.

As a solution, scholars-built Grad-CAM (Gradient-weighted Class Activation Mapping), which clarifies by showing the areas of an image that guided the model's choice. However, these tools are not always reliable in building complete trust. Interpretability is growing, so explanations should be true and relevant to clinical practice for maximum utility.

Still, the idea of ethics is increasingly important. What should hospitals do when AI misses a diagnosis? Who is the law responsible when medical predictions go wrong, and a patient is affected? Since these situations are complicated, we need strong rules and guidelines to solve them.

A further challenge is called bias. If the data used for training comes only from certain groups—patients from a specific country, older or younger groups, or those with particular socioeconomic status—the model is likely to fail when it comes to underrepresented groups. Influencing the medication a patient receives can sometimes worsen existing health differences.

Some privacy concerns appear when data is exchanged for building and training models. Large and varied datasets from various hospitals or districts are regularly needed for multi-modal CNNs. All this information must be de-identified and safe to maintain trust and obey regulations.

11. Future Directions

11.1. Incorporating More Imaging Modalities

New possibilities for future improvement include adding other imaging methods besides CT, X-ray, and ultrasound to multi-modal CNN frameworks. Several useful findings can be produced when imaging data from MRI and PET is applied, improving diagnosis not just for COVID-19 but for many other illnesses, including brain diseases, heart and cancer.

Back when the pandemic began, doctors did not use CT scans to diagnose COVID-19, but they have allowed experts to spot how the disease leaves lasting marks in the lungs. Using brain MRIs, doctors can now investigate how COVID-19 may cause encephalopathy and stroke. Like heart scans, cardiac MRIs are able to spot myocarditis which has turned up in people who have tested positive for COVID. With their help, MRI images can teach us more about how COVID-19 affects the body's systems.

Since PET shows how metabolism occurs in real-time, it may become increasingly valuable in future multi-modal strategies. PET scans could help identify ways the disease changes the blood vessels or measure very early inflammation, revealing information about its development at the cell level. If PET and CT or MRI are used, the resulting inputs can let CNN diagnose infection and damage the whole body.

Combining these different images is hard since some are low resolution, some are high resolution, the data is coded in multiple ways, and they are designed to focus on different anatomy. Still, new ideas in cross-modal and transfer learning are starting to join these worlds. Architectures built on CNNs are designed now to work with heterogeneous information by separating the inputs into distinct branches and using fusion strategies to merge the outputs.

At the same time, adding these new modes into practice also makes us rethink how available and standardized the data is. Fewer and harder-to-access MRI and PET datasets exist than X-ray and CT scans. For this reason, it is necessary to create a common database for these new techniques, accurately annotated with clinical and outcome details.

Integrating MRI and PET imaging into multi-modal CNN models will allow us to precisely diagnose forecast long-term disorders, determine patient outcomes, and judge treatment responses. Diagnostic models may become fully functional clinical decision support systems, opening the way for even better precision medicine.

11.2. Federated Learning for Multi-Center Collaboration

A key development for this type of network is using federated learning, which helps institutions collaborate to design powerful artificial intelligence models using their data instead of sharing it. Because of this approach, medical imaging and COVID-19 diagnosis can greatly change how we train and deploy CNNs in various healthcare centers worldwide.

Federated learning helps hospitals or research centers build one global model on their local information. The model's weight updates are sent only from learned clients to a server coordinating updates to the international model. All the raw data, including scans and private information, stays in the hospital. Because of this approach, private data is always secure, which gives healthcare organizations the confidence to collaborate with other institutions.

Hospitals from different specialties in federated learning can add their modalities to a joint multi-modal CNN. Some places give excellent CT services, others specialize in X-rays, and various hospitals may center on MRIs of the heart or brain. Because of this variety, the resultant model is better able to function equally well for patients from different locations.

Federated learning tackles the problem of unequal amounts of data and inequitable portrayals of users. There is often the tendency for centralized data to reflect more examples of some demographics or stages of a disease. When we compare, data used in federated learning come from various places, cultures, and treatment centers, helping keep the model fair and stopping it from only working in a particular area.

Bringing federated learning into practice is challenging from a technological perspective. It demands that updates happen together on various systems, that strong encryption is in place, and that there are ways to process all kinds of data from different sites. Studying the issues of communication usage and model convergence is a busy field under this paradigm.

Even so, thanks to TensorFlow Federated and PySyft, as well as other frameworks, using these models is possible. Early studies have proven that federated learning can teach diagnostic models to work with COVID-19, diabetic retinopathy, and breast cancer.

In the future, federated learning could replace other approaches as the regular way to make multi-modal CNNs for healthcare. Blockchain allows for the best benefits of big data and the storage of money locally. As a result, truly joint AI models that meet clinical standards, are safe and support worldwide use can be created.

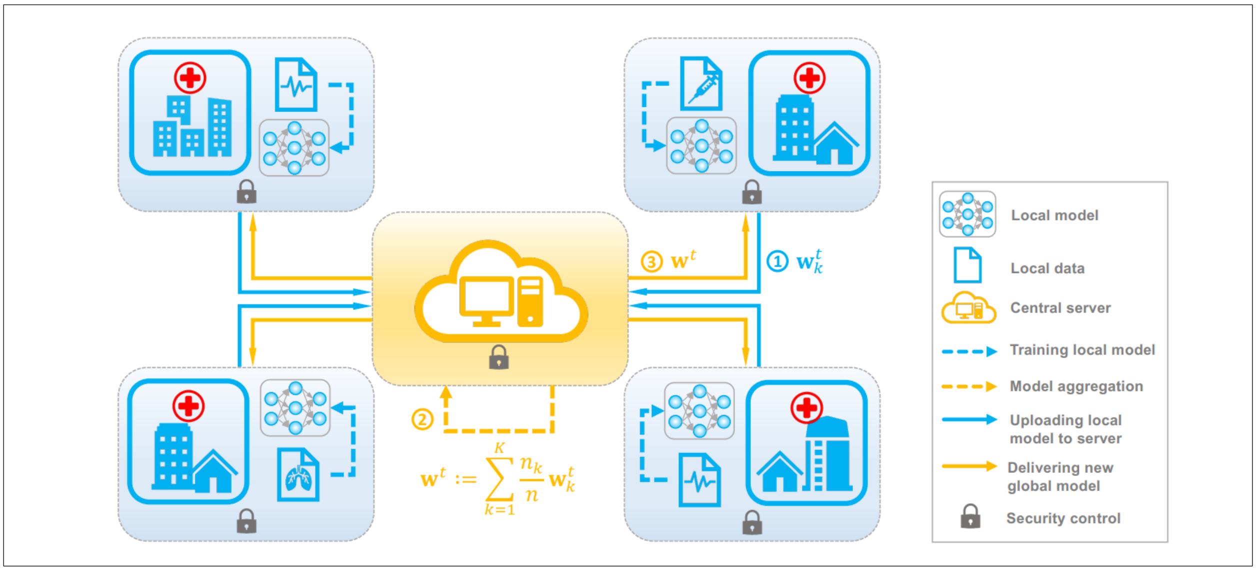

Figure 3 Data Flow in Federated Learning

12. Conclusion

Summary of Key Takeaways

COVID-19 diagnosis in healthcare has benefited from introducing multi-modal CNN frameworks as one of its leading innovations. We described how the pandemic made it necessary for new, improved diagnostic methods and showed how Convolutional Neural Networks stepped up to the task.

Our team highlighted that breezing scans with just one approach, standard X-rays or CT images, did not always show the bigger picture of a person's illness. Because of this, multi-modal approaches began to include X-ray, CT, and ultrasound as complementary parts of one strategy.

CNNs are the leading choice for this innovation because their skills in processing data make it easier to discover complex patterns in scans. Thanks to their organization, neural networks can spot hints of COVID-19 that ordinary people could overlook.

We discussed how these models were created, learned from data, and optimized—starting with getting the data and preprocessing it, designing it, training it, testing its results, and using what was learned. Sufficient importance was assigned to reliable data augmentation, carefully chosen hyperparameters, and evaluation methods using accuracy, precision, recall, F1-score, and AUC.

Various tests and comparisons showed that multi-modality cancer detection systems are better and more reliable than traditional one-type methods for diagnosing cancer quickly and flexibly. We ensured CNNs could function smoothly in hospitals so that tasks could be handled immediately, triage could happen instantly, and choices could be made as soon as possible.

Naturally, growing your business presents several obstacles. Like exploring any other field, data scientists in this area must address data scarcity, problems with how data is annotated, the interpretability of their models, and ethical issues. Still, putting the right guardrails in place, explanatory methods, and group work can help overcome these barriers.

After that, we realized the growing importance of imaging methods like MRI and PET. We focused on how federated learning could let us share data to train global models while keeping everything private.

Final Thoughts on CNN and Multi-Modal Integration

Blending AI and medical imaging, represented by multi-modal CNN frameworks, is more important than just a new technology—it changes how we handle healthcare. Originally meant to help in a crisis, this solution is now ready to stay in diagnostic medicine, raising accuracy and efficiency and making diagnoses more personalized.

CNNs are now able to match or go beyond how people diagnose images. Still, if given access to data from several sources, they become even wiser about making decisions, just as an experienced doctor does. Being able to link various imaging results into a single, dependable analysis is what gives multi-modal CNNs their strength.

The results will be important in many other circumstances. As the models improve and change, they are planned for use in identifying cancer, stroke, chronic lung disease, and many other neurological problems. As explainable AI and ethical guidelines develop further, they won't take the role of doctors but will greatly enhance the care they give to patients.

Multi-modal CNNs stand out because they combine technological, medical, and data expertise to boost our understanding rather than take our place. They use technology to show how new methods of diagnosis will be swifter, more precise, fairer, and easier to scale globally.

Compliance with ethical standards

Disclosure of conflict of interest

No conflict of interest to be disclosed.

References

[1] Abbas, A., Abdelsamea, M. M., and Gaber, M. M. (2020). Classification of COVID-19 in chest X-ray images using DeTraC deep convolutional neural network. arXiv preprint arXiv:2003.13815. https://arxiv.org/abs/2003.13815

[2] Afshin, S., Khodatars, M., Jafari, M., Ghassemi, N., Sadeghi, D., Moridian, P., ... and Gorriz, J. M. (2020). Automated detection and forecasting of COVID-19 using deep learning techniques: A review. arXiv preprint arXiv:2007.10785. https://arxiv.org/abs/2007.10785

[3] Harmon, S. A., Sanford, T. H., Xu, S., Turkbey, E. B., Roth, H., Xu, Z., ... and Wood, B. J. (2020). Artificial intelligence for the detection of COVID-19 pneumonia on chest CT using multinational datasets. Nature Communications, 11(1), 4080. https://doi.org/10.1038/s41467-020-17971-2

[4] Basu, S., Mitra, S., and Saha, N. (2020). Deep Learning for Screening COVID-19 using Chest X-Ray Images. https://doi.org/10.1101/2020.05.04.20090423

[5] Hofmanninger, J., Prayer, F., Pan, J., Rohrich, S., Prosch, H., and Langs, G. (2020). Automatic lung segmentation in routine imaging is primarily a data diversity problem, not a methodology problem. European Radiology Experimental, 4(1), 50. https://doi.org/10.1186/s41747-020-00173-2

[6] Hu, S., Gao, Y., Niu, Z., Jiang, Y., Li, L., Xiao, X., ... and Yang, G. (2020). Weakly supervised deep learning for COVID-19 infection detection and classification from CT images. arXiv preprint arXiv:2004.06689. https://arxiv.org/abs/2004.06689

[7] Hussain, E., Hasan, M., Rahman, A., Lee, I., Tamanna, T., and Parvez, M. Z. (2020). CoroDet: A deep learning-based classification for COVID-19 detection using chest X-ray images. Chaos, Solitons and Fractals, 142, 110495. https://doi.org/10.1016/j.chaos.2020.110495

[8] Islam, M. Z., Islam, M. M., and Asraf, A. (2020). A combined deep CNN-LSTM network for the detection of novel coronavirus (COVID-19) using X-ray images. Informatics in Medicine Unlocked, 20, 100412. https://doi.org/10.1016/j.imu.2020.100412

[9] Jin, C., Chen, W., Cao, Y., Xu, Z., Tan, Z., Zhang, X., ... and Feng, J. (2020). Development and evaluation of an artificial intelligence system for COVID-19 diagnosis. Nature Communications, 11(1), 5088. https://doi.org/10.1038/s41467-020-18685-1

[10] Khan, A. I., Shah, J. L., and Bhat, M. M. (2020). CoroNet: A deep neural network for detection and diagnosis of COVID-19 from chest x-ray images. Computer Methods and Programs in Biomedicine, 196, 105581. https://doi.org/10.1016/j.cmpb.2020.105581

[11] Oh, Y., Park, S., and Ye, J. C. (2020). Deep Learning COVID-19 Features on CXR using Limited Training Data Sets. arXiv.org. https://arxiv.org/abs/2004.05758

[12] Waheed, A., Goyal, M., Gupta, D., Khanna, A., Al-Turjman, F., and Pinheiro, P. R. (2020). CovidGAN: Data augmentation using auxiliary classifier GAN for improved COVID-19 detection. IEEE Access, 8, 91916–91923. https://doi.org/10.1109/ACCESS.2020.2994762

[13] Wang, L., Lin, Z. Q., and Wong, A. (2020). COVID-Net: A tailored deep convolutional neural network design for detection of COVID-19 cases from chest X-ray images. Scientific Reports, 10(1), 19549. https://doi.org/10.1038/s41598-020-76550-z

[14] Wang, X., Deng, X., Fu, Q., Zhou, Q., Feng, J., Ma, H., ... and Wang, B. (2020). A weakly-supervised framework for COVID-19 classification and lesion localization from chest CT. IEEE Transactions on Medical Imaging, 39(8), 2615–2625. https://doi.org/10.1109/TMI.2020.2995965

[15] Prior-Attention Residual learning for more discriminative COVID-19 screening in CT images. (2020). IEEE Journals and Magazine | IEEE Xplore. https://ieeexplore.ieee.org/document/9094362

[16] Kaushik, P., and Jain, M. (2018). Design of low power CMOS low pass filter for biomedical application. International Journal of Electrical Engineering and Technology (IJEET), 9(5).

[17] Kumar, Y., Saini, S., and Payal, R. (2020). Comparative Analysis for Fraud Detection Using Logistic Regression, Random Forest and Support Vector Machine. SSRN Electronic Journal.

[18] Höppner, S., Baesens, B., Verbeke, W., and Verdonck, T. (2020). Instance-Dependent Cost-Sensitive Learning for Detecting Transfer Fraud. arXiv preprint arXiv:2005.02488.

[19] Niu, X., Wang, L., and Yang, X. (2019). A Comparison Study of Credit Card Fraud Detection: Supervised versus Unsupervised. arXiv preprint arXiv:1904.10604.

[20] Bhat, N. (2019). Fraud detection: Feature selection-over sampling. Kaggle. Retrieved from https://www.kaggle.com/code/nareshbhat/fraud-detection-feature-selection-over-sampling

[21] Olaitan, V. O. (2020). Feature-based selection technique for credit card fraud detection. Master's Thesis, National College of Ireland. Retrieved from https://norma.ncirl.ie/5122/1/olaitanvictoriaolanlokun.pdf

[22] Dal Pozzolo, A., Boracchi, G., Caelen, O., Alippi, C., and Bontempi, G. (2017). Credit card fraud detection: A realistic modeling and a novel learning strategy. IEEE Transactions on Neural Networks and Learning Systems, 29(8), 3784–3797. https://doi.org/10.1109/TNNLS.2017.2736643

[23] Carcillo, F., Dal Pozzolo, A., Le Borgne, Y. A., Caelen, O., Mazzer, Y., and Bontempi, G. (2019). Scarff: A scalable framework for streaming credit card fraud detection with spark. Information Fusion, 41, 182–194. https://doi.org/10.1016/j.inffus.2017.09.005

[24] West, J., and Bhattacharya, M. (2016). Intelligent financial fraud detection: A comprehensive review. Computers and Security, 57, 47–66. https://doi.org/10.1016/j.cose.2015.09.005

[25] Zareapoor, M., and Shamsolmoali, P. (2015). Application of credit card fraud detection: Based on bagging ensemble classifier. Procedia Computer Science, 48, 679–685. https://doi.org/10.1016/j.procs.2015.04.201

[26] Bhattacharyya, S., Jha, S., Tharakunnel, K., and Westland, J. C. (2011). Data mining for credit card fraud: A comparative study. Decision Support Systems, 50(3), 602–613. https://doi.org/10.1016/j.dss.2010.08.008

[27] Patel, H., and Zaveri, M. (2011). Credit card fraud detection using neural network. International Journal of Innovative Research in Computer and Communication Engineering, 1(2), 1–6. https://www.ijircce.com/upload/2011/october/1_Credit.pdf

[28] Puneet Kaushik, Mohit Jain , Gayatri Patidar, Paradayil Rhea Eapen, Chandra Prabha Sharma (2018). Smart Floor Cleaning Robot Using Android. International Journal of Electronics Engineering. https://www.csjournals.com/IJEE/PDF10-2/64.%20Puneet.pdf

[29] Duman, E., and Ozcelik, M. H. (2011). Detecting credit card fraud by genetic algorithm and scatter search. Expert Systems with Applications, 38(10), 13057–13063. https://doi.org/10.1016/j.eswa.2011.04.102

[30] Kaushik, P., Jain, M., and Jain, A. (2018). A pixel-based digital medical images protection using genetic algorithm. International Journal of Electronics and Communication Engineering, 31-37. http://www.irphouse.com/ijece18/ijecev11n1_05.pdf

[31] Kaushik, P., Jain, M., and Shah, A. (2018). A Low Power Low Voltage CMOS Based Operational Transconductance Amplifier for Biomedical Application. https://ijsetr.com/uploads/136245IJSETR17012-283.pdf

[32] Kaushik, P., and Jain, M. (2018). Design of low power CMOS low pass filter for biomedical application. International Journal of Electrical Engineering and Technology (IJEET), 9(5).

[33] Nabati, R., and Qi, H. (2019). "RRPN: Radar Region Proposal Network for Object Detection in Autonomous Vehicles." 2019 IEEE International Conference on Image Processing (ICIP), Taipei, Taiwan, 2019, pp. 3093-3097, doi: 10.1109/ICIP.2019.8803392.

[34] Rawat, W., and Wang, Z. (2017). "Deep Convolutional Neural Networks for Image Classification: A Comprehensive Review." Neural Computation, 29(9), pp. 2352-2449, Sept. 2017, doi: 10.1162/neco_a_00990.

[35] Wang, W., et al. (2019). Medical image classification using deep learning. In Intelligent Systems Reference Library (pp. 33–51). https://doi.org/10.1007/978-3-030-32606-7_3

[36] Alom, M. Z., et al. (2018). The history began from AlexNet: A comprehensive survey on deep learning approaches. arXiv. https://arxiv.org/abs/1803.01164

[37] Frid-Adar, M., et al. (2018). GAN-based synthetic medical image augmentation for increased CNN performance in liver lesion classification. Neurocomputing, 321, 321–331. https://doi.org/10.1016/j.neucom.2018.09.013

[38] Jiao, L., and Zhao, J. (2019). A survey on the new generation of deep learning in image processing. IEEE Access, 7, 172231–172263. https://doi.org/10.1109/ACCESS.2019.2956508

[39] Wang, L., Chen, W., Yang, W., Bi, F., and Yu, F. R. (2020). A state-of-the-art review on image synthesis with generative adversarial networks. IEEE Access, 8, 63514–63537. https://doi.org/10.1109/ACCESS.2020.2982224

[40] Shorten, C., and Khoshgoftaar, T. M. (2019). A survey on image data augmentation for deep learning. Journal of Big Data, 6(1), Article 60. https://doi.org/10.1186/s40537-019-0197-0

[41] Kayalibay, B., et al. (2017). CNN-based segmentation of medical imaging data. arXiv. https://arxiv.org/abs/1701.03056

[42] Kaushik, P., and Jain, M. (2018). A low power SRAM cell for high speed applications using 90nm technology. International Journal of Electrical Engineering, 10(2), 6. https://www.csjournals.com/IJEE/PDF10-2/66.%20Puneet.pdf

[43] Esfahani, S. N., and Latifi, S. (2019, July 13). A survey of state-of-the-art GAN-based approaches to image synthesis. In Proceedings of the 9th International Conference on Computer Science, Engineering and Applications (CCSEA 2019). https://doi.org/10.5121/csit.2019.90906

[44] Yang, S., Jiang, L., Cao, Z., Wang, L., Cao, J., Feng, R., Zhang, Z., Xue, X., Shi, Y., and Shan, F. (2020). Deep learning for detecting corona virus disease 2019 (COVID-19) on high-resolution computed tomography: a pilot study. Annals of Translational Medicine, 8(7), 450. https://doi.org/10.21037/atm.2020.03.132

[45] Xu, J., Glicksberg, B. S., Su, C., Walker, P., Bian, J., and Wang, F. (2019). Federated Learning for Healthcare Informatics. ArXiv.org. https://arxiv.org/abs/1911.06270

[46] Yildirim, M., Eroğlu, O., Eroğlu, Y., Çinar, A., and Cengil, E. (2022). COVID-19 Detection on Chest X-ray Images with the Proposed Model Using Artificial Intelligence and Classifiers. New Generation Computing, 40(4), 1077–1091. https://doi.org/10.1007/s00354-022-00172-4

[47] Kaushik, P. (2018). STUDY AND ANALYSIS OF IMAGE ENCRYPTION ALGORITHM BASED ON ARNOLD TRANSFORMATION. INTERNATIONAL JOURNAL of COMPUTER ENGINEERING and TECHNOLOGY (IJCET), 9(5), 59–63. https://iaeme.com/Home/article_id/IJCET_09_05_008