[1]S. Deepika,

Dr. M. Rajeswari

Evidence-Efficient Compressed Deep Ensemble Resnet for COVID-19 Detection

JES Journal of Electrical Systems

Abstract: - The new coronavirus illness 2019 (COVID-19) put enormous strain on healthcare systems globally. Early diagnosis of COVID-19 is critical for controlling the spread of the COVID-19 pandemic and relieving pressure on health-care systems. Deep Learning (DL) models have a significant role in the COVID-19 detection by adopting the medical imaging techniques like Computerized Tomography (CT) and Chest X-ray (CXR). Many DL models have been developed to detect and diagnose the COVID-19 cases at earlier stages. Amongst, Hybrid Deep Transfer Learning (HDTL) model is developed using an Enhanced Convolution Restricted Boltzmann Machine (ECRBM) and ResNet model to improve the Covid-19 recognition tasks. But still, this model results in epistemic uncertainty which effects the performance of DL models employed for COVID-19 detection and diagnosis. On considering this, an effective DL framework is proposed to solve the challenges of epistemic uncertainty in DL models to provide a rapid detection of various diseases using CT and CXR images. This model considers the uncertainty issue from the theory of evidence perspective. The softmax layer of ResNet models is modified with Dirichlet density parameters to represent learner predictions as a distribution over possible softmax outputs. The model has a specific loss function which is minimized through standard backprop and fitted to data by minimizing the Bayes risk with respect to the L2-Norm loss. The resultant predictor is a Dirichlet distribution on class probabilities which provides more detailed uncertainty model than the point estimate of the softmax-output deep nets (evidence predictor). The developed model results in increased accuracy and estimated uncertainty; however, linearly enhancing the size makes the deep ensemble unfeasible for memory-intensive tasks. To address the problem of memory intensive tasks, the model pruning and quantization techniques are performed along with deep ensemble and analyzed the effect in the context of uncertainty metrics. The magnitude-based weight pruning is used to prune the ResNet model, until it reached the required sparsity and a few epochs retrain the pruned networks. It is noticeable that increased disagreement in deep ensemble models will implies increased uncertainty which helps in making more robust predictions. The complete work is termed as Evidence-Efficient Compressed Deep Ensemble (EECDE) – ResNet which is appropriate for memory-intensive and uncertainty aware tasks. The test result show that the proposed model achieves accuracy of 98.29 % and 97.69% on CT and CXR mages for COVID-19 identification.

Keywords: COVID-19, Compuerized Tomography, Chest X-ray, Dirichlet distribution, Deep Ensemble.

1. INTRODUCTION

At the end of 2019, a new virus spread through China and caused a major pandemic [1,2]. This virus is also called as COVID-19 which can cause a number of different symptoms leading to chronic respiratory disease and acute respiratory distress syndrome, and these can be deadly [3,4]. Due to the rapid propagation of this virus, the COVID-19 had been found in more than 350 million people around the world within a consecutive days. And, it has also killed more than 5 million people directly [5]. The World Health Organization (WHO) declares a state of emergency in 2020 due to the likelihood of fast transmission of COVID-19. So, it is important for the physician to detect the COVID-19 at its earlier stages. The definitive predictive tool for diagnosing COVID-19 is the Reverse Transcription Polymerase Chain Reaction (RT-PCR), but it is insufficiently sensitive [6, 7]. It cannot accommodate the daily increase in patient demand. The procedure is not only time-consuming, but it can also perform some wrong calculations in an emergency conditions.

Medical Imaging systems like CT and CXR is one of the most important fields of study and plays a vital role in identifying and diagnosing variety of diseases, including COVID-19. Many investigations and more attention are paid to imaging modalities for the efficient prediction tasks. Therefore, interpretation of these images requires expertise and necessitates several algorithms to enhance, accelerate and make an accurate diagnosis [8, 9]. In recent times, researchers are keenly interested in the topic of artificial intelligence (AI) techniques that assist

[1]Ph.D. Research Scholar
[2]Assistant Professor
Department of Computer Science,
PSGR Krishnammal College for Women, Coimbatore, India

doctors in diagnosing many diseases, especially developing technological capabilities of detecting Covid-19 virus [10].

AI is of two modes, Machine Learning (DL) and Deep Learning (DL), both the methods plays a vital role in the earlier detection and diagnosis of COVID-19 and lung diseases. In addition, ML and DL methods have become popular in medical image processing and diagnostic problems to identify different kinds of diseases [11]. The ML models like Support Vector Machine (SVM) [12], k-Nearest Neighbour (k-NN) [13], Artificial Neural Network (ANN) [14] and so on., are used as decision-making tools which are highly helpful in detecting and diagnosing the COVID-19 cases as early as possible. On the other hand, ML classifier's performances are decided based on the predefined parameter alone. Also, the training period is lengthy and inappropriate for large volumes of data.

DL is a emerging field which is applied to diagnose a variety of other chronic conditions especially COVID-19, and to enable doctors in determining medical decisions. This method enables immediate disease detection for analysing complex medical images [15]. According to this network structure, DL approaches are often classed as Convolutional Neural Network (CNN), Recurrent Neural Network (RNN), Transfer Learning (TL) and so on. Among these models, CNN is a popular DL method which aims to train clinical image structures and attributes for diagnosing all forms of disease using enhanced clinical images [16]. Many investigations on DL based medical imaging detections are developed for COVID-19 for the early diagnosis and detection.

Different pre-trained CNN structures like InceptionResNet V2, InceptionResNetV3, VGGNet, DenseNet 121, ResNet-50 are available to utilized in solving a problem by quickly learning the appropriate parameter. For instances, a new TL-based uncertainty-aware technique [17] was developed for the effective detection of COVID-19 patients from and CT and CXR scanned images. This method extracts important and distinctive features from and CT and CXR images in a systematic way using different qualitative pre-trained CNN models. In which, ResNet50 provided better results for COVID-19 detection. However, this technique performed better only for selecting appropriate layer for ResNet architecture. An appropriate layers selected manually for given input image is highly challengeable.

The fore-presented issues are resolved by developing an ESMO algorithm is used to select appropriate layers of ResNet architecture [18]. In this model, Spider Monkey (SM) selects a number of layers and parameter to fine-tune the architecture. Categorical Cross-Entropy Loss (CCEL) (fitness) is computed to find the best optimal solutions. The feature extracted from ESMO-ResNet is trained and classified images as normal or COVID-19 using Softmax classifier. However, there are constantly some content differences between data sets during DTL classifier which severely influence the identification performances classifier.

On addressing this, Hybrid Deep Transfer Learning (HDTL) model is developed based on Enhanced Convolution Restricted Boltzmann Machine (ECRBM) and ResNet model to improve Covid-19 recognition tasks [19]. The ResNet structure is pre-trained on ImageNet and transferred to the Covid-19 training samples as a small target set. Then, the Contrastive Divergence (CD) is used to fine-tune the parameters of softmax regression and each CRBM layer to obtain a better trained DTL classifier. The test samples in the target set will be used as the input information, and the final identification result will be determined from the newly trained model. However, this model results in epistemic uncertainty which effects the performance of DL models employed for COVID-19 detection and diagnosis.

On considering this, an Energy-Efficient Compressed Deep Ensemble (EECDE) – ResNet model is proposed in this paper to solve the challenges of epistemic uncertainty in DL models for the earlier detection of COVID-19 using CT and CXR images. This model considers the uncertainty issue from the theory of evidence perspective. The softmax layer of ResNet models is modified with Dirichlet density parameters to represent learner predictions as a distribution over possible softmax outputs. The model has a specific loss function which is minimized through standard backprop and fitted to data (CT and CXR) by minimizing the Bayes risk with respect to the L2-Norm loss. The resultant predictor is a Dirichlet distribution on class probabilities which provides more detailed uncertainty model than the point estimate of the evidence predictor (softmax). The developed model results in increased accuracy and estimated uncertainty; however, linearly enhancing the size makes the deep ensemble unfeasible for memory-intensive tasks.

To address the problem of memory intensive tasks, the model pruning and quantization techniques are performed along with deep ensemble and analyzed the effect in the context of uncertainty metrics. The magnitude-based weight pruning is used to prune the ResNet structure, until the model reached the required sparsity and a few epochs retrain the pruned networks. Finally, each model's weights are quantized from higher level bit representation to lower level representation. Thus the proposed model could obtain a lightweight ensemble by applying pruning and quantization, which is robust to unknown instances during the inference time. It is noticeable that increased disagreement in deep ensemble models will implies increased uncertainty which helps in making more robust predictions of COVID-19.

The remaining section of this paper is structured as follows: Section II presents the related research performed on predicting and classifying COVID-19. The proposed research paradigm is explained in Section III, and its effectiveness is shown in Section IV. Conclusions and suggestions for further research are presented in Section V.

2. LITERTURE SURVEY

Biradar et al. [20] introduced an effective DL model for COVID-19 infection detection in lungs using pretrained CNN models. This model includes the design of user-friendly web and mobile application interfaces for uploading CXR images and monitoring the general health parameters of the patient. The COVID-19 images are collected and augments using translation, flipping and contrast change to enhance the dataset to solve both the issues of limited dataset and the imbalanced dataset. Then, the DL model was applied for the classification tasks. Moreover, CovIdentify mobile application was developed which supports the users seamlessly access the services and help them to identify COVID positively or negatively. However, this model was expensive and have high-end work tasks.

Deb et al. [22] introduced a multi-model ensemble based deep CNN (MME-DCNN) structure for the identification of COVID-19. Initially, the individual DCNN structure was used for feature extraction. Then, an ensemble of pre-trained DCNN models extracts the low level characteristics from the CXR images before transferring them to a fully connected layer for classification. However, the accuracy might be declined because it was only examined on a modest number of X-ray samples.

Huang et al. [23] constructed a lightweight CNN-based network on COVID-19 detection using CXR and CT images. Initially, the collected images was pre-processed to normalize the pixel values and unify the image format. Then, various TL models were fine-tuned and LightEfficientNetV2 model was used to optimize the training spped and parameter efficiency. Also, LightEfficientNetV2 utilizes a non-uniform scaling strategy, which can increase the number of layers in more subsequent layers. Moreover, Gridsearch Cross Validation (CV) was employed to explore the parameter efficiency in CNN model. However, this model lacks efficiency while training the larger data samples.

Kogilavani et al. [24] presented a DL model for the COVID-19 detection based on lung CT scan images. In this process, collected dataset was pre-processed to make all images with same size and it was splitted up into training, validation, and test data. The training dataset was fed into different CNN architectures and was evaluated using a validation dataset with different epochs. Finally, the models were being tested by supplying test data. However, this method was acquired with low convergence rate.

Zouch et al. [25] constructed a DL models to detect the COVID-19 from CT and CXR images. In this model, three augmentation strategies like random rotation, translation and horizontal flip was applied on the collected dataset. Then, the augmented data were given as input to the two pre-trained CNN models like VGGNet and ResNet models for the detection of COVID -19. However, considerable period of time was required to train the data.

Gour & Jain [26] presented an automated COVID-19 detection from CXR and CT images with stacked ensemble CNN (SE-CNN). In this model different sub-models have been obtained from the VGG19 and the Xception models during the training. Then, obtained sub-models are stacked together using softmax classifier. The stacked CNN model combines the discriminating power of the different CNN's sub-models and detects COVID-19 from the radiological images. However, this model needs more effective parameters to diagnose the COVID-19.

Kathamuthu et al. [27] developed a deep transfer learning-based CNN (DTL-CNN) model using CT scan images for COVID-19 detection and diagnosis. At first, the Images from CT scans of the lungs were collected and pre-processed. The pre-processed images were given as input to pre-trained CNN models to develop a binary classifier for Chest CT scans and predict the presence of COVID-19. Next, the top layers of the pre-trained model was customized and constructive model was developed for the final detection tasks. However, this model results in higher generalization error.

Ahmed et al. [28] constructed lightweight CNN (LW-CovidNet) method to automatically detect covid-19 infected regions from CXR. In this model, standard and depth-wise separable CNN were integrated to aggregate the high level features from the collected dataset. Then, the information loss was compensated by enhancing the receptive field of this model. The detection boundaries of disease regions representations were then enhanced by edge-attention method by applying heat maps which forecasts for efficient COVID-19 detection. However, the performance was lower on smaller datasets.

Bhattacharjee et al. [29] developed a DeepCOVNet Model for COVID-19 detection using CXR images. The DeepCOVNet was a modified version of CNN which composed of three Conv2D layers, each followed by a Maxpool2D layer, a 25% dropout layer, and two fully connected (FC) layer. Initially, the CXR images of COVID-19 patients was collected. Then, the collected images were pre-processed. The pre-processed images were given as input to different pre-trained CNN and DeepCOVNet model for the COVID-19 prediction tasks. However, the training epoch was high which leads to high over-fitting issues.

Foysal et al. [30] presented a Deformable DCNN (DDCNN) for the COVID-19 identification tasks from Chest CT images. In this model, two different DCNN model like conventional CNN and the ResNet-50 for COVID-19 detection from chest CT with a strategic emphasis on finding the impact of the deformable concept through a comparative performance analysis among the normal and deformable forms. A fifteen-layered DCNN model was developed and then its deformable form was created. Deformable form of normal CNN model was performed by replacing two convolution layers with deformable convolution layers. In addition, gradient class activation mapping (Grad-CAM) technique was applied to visualize and check the targeted regions for localizing effort at the final convolutional layer for the robust COVID-19 prediction tasks. However, this approach had a low convergence rate.

3. PROPOSED METHODOLOGY

In this section, the EECDE – ResNet model is developed to resolve the challenges of uncertainty and memory intensive for the COVID-19 detection. The figure 1 defines the block diagram of proposed model.

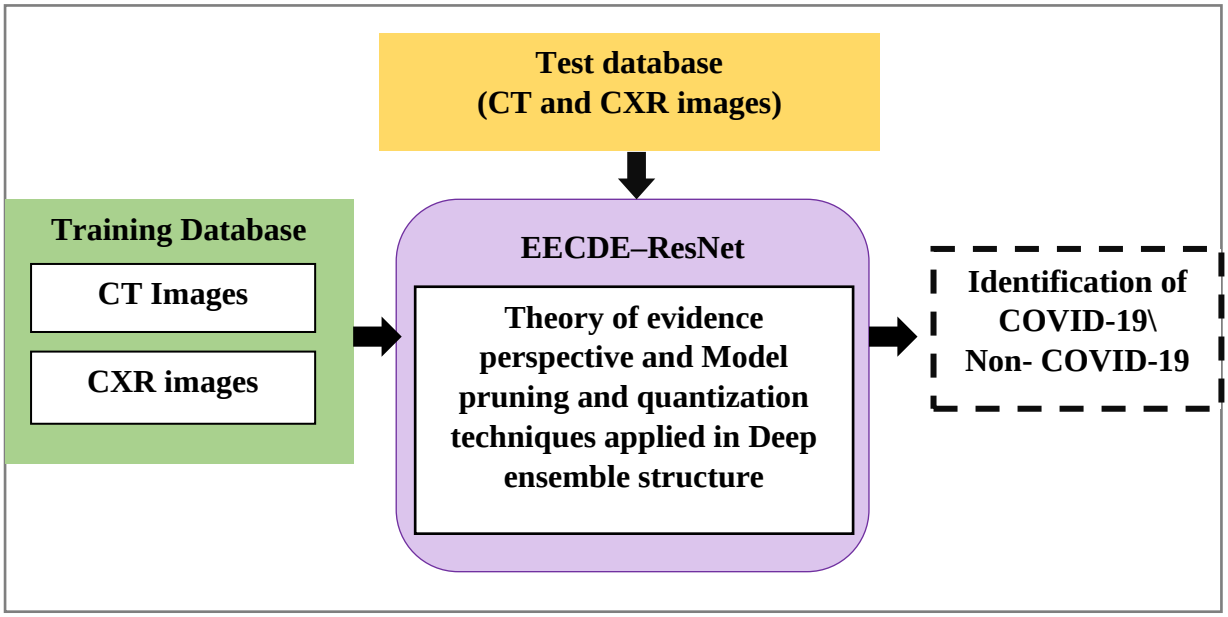


Figure 1. Block Diagram of EECDE–ResNet System

3.1 Theory of Evidence perspective for uncertainty estimation problem

A Theory of Evidence approach [31, 32] will be used to tackle the uncertainty estimation problem. Here, the typical result of a softmax classification network is considered as the parameters of a discrete distribution. When the parameters of a Dirichlet density are substituted for these values, the learner's predictions are no longer a point estimate, but rather a distribution across all potential softmax outputs. This density can be represented as a production line for these point estimates. The generated model has a specific loss function, which is minimized subject to neural net weights using standard backpropagation model. The proposed model adopts the large margin on two applications while modifying the high-quality uncertainity is an essential tasks. Particularly, the prediction distribution of these models should be closer to the maximum entropy settings, when supplied with input from a distribution different from that of the training samples (i.e., CT and CXR pictures). The softmax is more resilient to disease detection tasks on CT and CXR image datasets since it keeps reporting high confidence for wrong classes for high rotations. All proposed model vectors are column vectors represented a u, with i^{th} element represented by the notation u_i. The symbol $\odot$ is used to indicate the Hadamard (element-wise) product.

3.3.1 Modeling Class Probabilities with Softmax and Its Limitations

In DL models, the softmax operator is utilized to convert the continuous activations of the output layer to class probabilities. The eventual model can be interpreted as a multinomial distribution whose parameters, hence discrete class probabilities, are determined by neural net outputs. For a I - class classification problem, the probability function for an observed tuple (u, v) will be

$$P(v|u,\theta) = Mult\left(v\middle|\sigma\big(g_1(u,\theta)\big), \dots\dots, \sigma\big(g_i(u,\theta)\big)\right) \quad (1)$$

In above Eq. (1), $Mult\ (\dots.)$ is a multinomial mass operation, $g_j(u,\theta)$ is the l^{th} resultant channel of an arbitrary neural net $g(.)$ and parameterized by θ, and $\sigma(x_l) = e^{x_l}/\sum_{i=1}^{I} e^{x_I}$ is the softmax function. While the continuous neural net is responsible for adjusting the ratio of class probabilities, softmax squashes these ratios into a simplex. The eventual softmax-squashed multinomial likelihood is then maximized with respect to the neural net parameters θ. The equivalent problem of minimizing the negative log-likelihood is preferred for computational convenience

$$-\ log\ p(v|u,\theta) = -\ log\ \sigma(g_v(x,\theta)) \quad (2)$$

The above-mentioned Eq. (2), is widely known as the cross-entropy loss. It is noteworthy that the probabilistic interpretation of the cross-entropy loss is minor Maximum Likelihood Estimation (MLE). ResNet estimates classification probabilities with the softmax function, which is notorious for inflating the probability of the predicted class. This can lead to unreliable uncertainty estimations, as the distance of the predicted label of a newly seen observation is not useful for the conclusion. ResNet estimates classification probabilities with the softmax function (evidence predictor), which is high for misclassified samples. However, this approach can accurately quantify uncertainty of its predictions.

3.1.2 Quantification of belief masses and uncertainty issues

The Dempster–Shafer Theory of Evidence (DST) is a generalization of the Bayesian theory to subjective probabilities. It assigns belief masses to subsets of a frame of discernment, which denotes the set of exclusive possible states, e.g., possible class labels for a sample. A belief mass can be assigned to any subset of the frame, including the whole frame itself, which represents the belief that the truth can be any of the possible states, e.g., any class label is equally likely or by assigning all belief masses to the whole frame as an opinion for the truth over possible states. Subjective Logic (SL) formalizes DST's notion of belief assignments over a frame of discernment as a Dirichlet Distribution. Hence, it allows one to use the principles of evidential theory to quantify belief masses and uncertainty through a well-defined theoretical framework. More specifically, SL considers a frame of K mutually exclusive singletons by enhancing a belief mass B_i for each singleton $i = 1, \dots, I$ and providing an overall uncertainty mass of x. These $I + 1$ mass values are all non-negative and sum up to one, i.e.,

$$x\ \sum_{i=1}^{I} B_i = 1 \quad (3)$$

In Eq. (3), $x \geq 0$ and $B_i \geq 0$ for $= 1, \ldots, I$. B_i for a singleton i is computed using the evidence for the singleton. Assume, $e_i \geq 0$ be the evidence derived for the kth singleton, then the belief B_i and the uncertainty x are computed in Eq. (5)

$$B_i = \frac{e_i}{T} \tag{4}$$

$$x = \frac{I}{T} \tag{5}$$

where $T = \sum_k^I (e_k + 1)$. It is considerable that the uncertainty is inversely proportional to the total evidence. When there is no evidence, the belief for each singleton is zero and the uncertainty is one. From the Bayesian modeling nomenclature, the term "evidence" is represented as a measure of the amount of support collected from data images in favor of a sample to be classified into a certain class. B_i assignment, i.e., subjective opinion wyhich corresponds to a Dirichlet distribution with parameters $\alpha_i = e_i + 1$. Whereas, a subjective opinion can be derived easily from the parameters of the corresponding Dirichlet distribution using $B_i = (\alpha_i - 1)\backslash T$, where $T = \sum_{k=1}^{I} a_k$ is referred to as the Dirichlet strength. The output of a ResNet classifier is a probability assignment over the possible classes for each sample. However, a Dirichlet distribution parametrized over evidence represents the density of each such probability assignment; hence it models second-order probabilities and uncertainty.

The Dirichlet distribution is a probability density function (pdf) for possible values of the probability mass operation (PMO). It is featured by I parameters, $\alpha = \alpha_1, \ldots\ldots, \alpha_I$ and is given by,

$$d(p\backslash\alpha) = \begin{cases} \frac{1}{\beta(\alpha)} \prod_{k=1}^{I} p_k^{a_k - 1} & for\ p \in T_I \\ 0 & otherwise \end{cases} \tag{6}$$

Where, T_I is the I-dimensional unit simplex,

$$T_I = \{p | \sum_{k=1}^{I} p_k = 1\ and\ 0 \leq p_1 \leq 1\} \tag{7}$$

In Eq. (7), $\beta(\alpha)$is the K-dimensional multinomial beta function.

Given an opinion, the expected probability for the i^{th} singleton is the mean of the corresponding Dirichlet distribution and computed as

$$\hat{p}_k = \frac{a_k}{T} \tag{8}$$

When an observation about a sample relates it to one of the I attributes, the corresponding Dirichlet parameter is incremented to update the Dirichlet distribution with the new observation. For instance, detection of a specific pattern on an image may contribute to its classification into a specific class. In this case, the Dirichlet parameter corresponding to this class should be incremented. This implies that the parameters of a Dirichlet distribution for the classification of a sample may account for the evidence for each class.

In this framework, it is noted that a neural network is capable of forming opinions for classification tasks as Dirichlet distributions. Assume, $\alpha_k = \langle \alpha_{k1}, \alpha_{k2}, \ldots\ldots, \alpha_{kI} \rangle$ is the parameters of a Dirichlet distribution for the classification of a sample k, then $(\alpha_{kl} - 1)$ is the total evidence estimated by the network for the assignment of the sample k to the l^{th} class. Furthermore, given these parameters, the epistemic uncertainty of the classification can easily be computed using Eq. (4) & (5).

3.1.3 Training to from Opinions

The softmax function provides a point estimate for the class probabilities of a sample and does not provide the associated uncertainty. On the other hand, multinomial opinions or equivalently
Dirichlet distributions can be used to model a probability distribution for the class probabilities.
Therefore, in this framework, a ResNet model is trained and utilized to form their multinomial opinions or the classification of a given sample k as a Dirichlet distribution $d(p_k|\alpha_k)$, where p_k is a simplex representing class assignment probabilities. The ResNet model for COVID-19 classification are very similar to simple pre trained CNN models. The only difference is that the softmax layer is replaced with an activation layer, e.g., ReLU, to ascertain non-negative output, which is taken as the evidence vector for the predicted Dirichlet distribution.

In a given sample k, assume $g(U_k|\Theta)$ depicts the evidence vector predicted by the network for the classification, where Θ is network parameters. Then, the corresponding Dirichlet distribution has parameters $\alpha_i = g"(U_k|\Theta) +$ 1. Once the parameters of this distribution is calculated, its mean, i.e., α_i/T_i, can be taken as an estimate of the class probabilities. Consider, V_k be a one-hot vector encoding the ground-truth class of observation u_k with $v_{kl} = 1$ and $v_{ki} = 0$ for all $\neq v$, and α_k be the parameters be the parameters of the Dirichlet density on the predictors. Initially, $d(p_k|\alpha_k)$ is treated as a prior on the probability $Mult(v_k|p_k)$ and obtain the negated logarithm of the marginal likelihood by integrating out the class probabilities and minimize with respect to the α_k parameters. This technique is well-known as the Type II Maximum Likelihood

$$\mathcal{L}_k(\Theta) = -\log\left(\int \prod_{l=1}^{I} p_{kl}^{y_{kl}} \frac{1}{\beta(\alpha_i)} \prod_{l=1}^{I} p_{kl}^{\alpha_{kl}-1} \, d(p_k)\right) = \sum_{l=1}^{I}(\log(\alpha_{kl})) \quad (9)$$

In another way, a loss function is computed using the Bayes risk with respect to the class predictor. For the cross-entropy, the bayes risk will read,

$$\mathcal{L}_k(\Theta) = \int[\sum_{l=1}^{I} -v_{kl}\log(p_{kl})]\frac{1}{\beta(\alpha_i)} \prod_{l=1}^{I} p_{kl}^{\alpha_{kl}-1} \, d(p_k) = \sum_{l=1}^{I}(\psi(T_k) - \psi(\alpha_{kl})) \quad (10)$$

In above Eq. (10), ψ is the digamma operation. The similar model can be applied also to the sum of squares loss $|v_k - p_k||$, resulting in

$$\mathcal{L}_k(\Theta) = \int ||v_k - p_k||_2^2 \frac{1}{\beta(\alpha_i)} \prod_{k=1}^{I} p_{kl}^{\alpha_{kl}-1} \, d(p_k) = \sum_{l=1}^{I} \mathbb{K}[v_{kl}^2 - 2v_{kl}p_{kl} + {p_{kl}}^2] = \sum_{l=1}^{I}(v_{kl}^2 - 2v_{kl}\,\mathbb{K}[p_{kl}] + \mathbb{K}[{p_{kl}}^2]) \quad (11)$$

The above observed the losses in Eq. (9) and (10) to generate excessively high belief masses for classes and exhibit relatively less stable performance than Eq. (11). The above loss aims to achieve the joint goals of minimizing the prediction error and the variance of the dirichlet experiment generated by the neural net specifically for each sample in the training set.

The loss over a batch of training samples can be computed by summing the loss for each sample in the batch. During training, the model may discover patterns in the data image and generate evidence for specific class labels based on these patterns to minimize the overall loss. However, when counter samples are observed during training, the parameters of the ResNet should be tuned by back propagation to generate smaller amounts of evidence for this pattern and minimize the loss of these samples. Unfortunately, when the number of counter-examples is limited, decreasing the magnitude of the generated evidence may increase the overall loss, even though it decreases the loss for the counter-examples. As a result, the ResNet model may generate some evidence for the incorrect labels. The Kullback-Leibler (KL) divergence term is incorporated into this loss function that regularizes proposed model's predictive distribution by penalizing those divergences from different labels that do not contribute to data fit. The loss with this regularizing term reads

$$\mathcal{L}(\Theta) = \sum_{k=1}^{n} \mathcal{L}_k(\Theta) + \gamma_s \sum_{k=1}^{n} KL[d(p_k|\alpha_{ki})| \, d(p_k|\langle 1, \ldots\ldots, \rangle)] \quad (12)$$

Where, $\gamma_s = \min\left(1.0, \frac{s}{10}\right) \in [0,1]$ is the annealing co-efficient, s is the index of the current training epoch, $d(p_k|\langle 1,..,1\rangle)$ is the uniform Dirichlet distribution, and lastly $\widetilde{\alpha_k} = y_k + (1 + y_k) \odot \alpha_k$ is the Dirichlet parameters after removal of the non-misleading evidence from predicted parameters α_k for sample k. The KL divergence term in the loss can be calculated as

$$KL[d((p_k|\breve{\alpha}_k)||d(p_k|1)] = \log\left(\frac{\Gamma(\sum_{i=1}^{I}\hat{\alpha}_{ki})}{\Gamma(I)\prod_{i=1}^{I}\Gamma(\hat{\alpha}_{ki})}\right) + (\sum_{l=1}^{I}(\hat{\alpha}_{kl} - 1))[\psi(\hat{\alpha}_{kl}) - \psi(\sum_{l=1}^{I}\hat{\alpha}_{kl})], \quad (13)$$

In above Eq. (13), 1 represents the parameter vector of I ones, $\Gamma(.)$ is the gamma function, $\psi(.)$ is the digamma function. By gradually increasing the effect of the KL divergence in the loss through the annealing coefficient, the ResNet model is allowed to explore the parameter space and avoid premature convergence to the uniform distribution for the misclassified samples, which may be correctly classified in the future epochs.

3.2 Model Compression Techniques on Uncertainty Estimation

The deep ensemble models improves accuracy through ensembling and effectively measures the uncertainty by maintaining diversity in the function space, thus increasing the disagreement in a distributional shift of inference

data. However, deep ensemble size linearly increase in parallel with the increase in ensemble members. The deep ensemble is not applicable in memory-intensive tasks because of its increased size.

In DL models, model compression techniques [33] dramatically reduce the model size compared with relatively insignificant to no accuracy degradation. Model pruning increases the model sparsity by zeroing the irrelevant parameters. Low-rank factorization decomposes the matrix tensor by estimating the informative parameters. Knowledge distillation learns a distilled model and trains a more compact network to reproduce a more extensive or ensemble network. Model quantization converts the parameter floating-point into 8 bits or less. In this framework, the deep ensemble size is reduced by using the model compression techniques and analyze the effect of model compression techniques on uncertainty estimation.

Assume, training image data is comprised of M data points $\{a, b\}$, where $x \in r^D$ represents $D-$dimensional features. For regression problems, the label $b \in r$ is a real-valued number. For classification problems, the label $b \in \{1, \ldots, I\}$ is assumed to be one of the I classes. The ResNet model is used to assign the prediction probability $p_\vartheta\ (b|a)$ over the labels, where ϑ denotes the ResNet's parameters. In this proposed model, an ensemble is developed which comprises Z number of independent neural networks.

3.2.1 Deep Ensemble

There are three basic uncertainty measurement formulas that represents the deep ensemble. The first step is to choose a valid scoring rule to use as a benchmark throughout training. The second method is to teach the network to function on its own. To further smooth the prediction distributions, they implemented adversarial training, however it was not as successful as the aforementioned two methods. If a predictive distribution is identical to the true distribution, then the scoring rule $T\ (p_\vartheta\ , w) \leq T\ (w, w)$ is correct. Here, p_ϑ stands for the predictive distribution and q stands for the genuine distribution. Training a network, then, entails finding the optimal value of $L(\vartheta)\ =\ -T\ (p_\vartheta\ , w)$. Key principles in deep ensembles include decoupled and randomly initialized DNN with common architecture. In the deep ensemble, all members are treated as a uniformly weighted mixture model, and the final prediction is the average of the predictions with Equation (14),

$$p(b|x) = Z^{-1} \sum_{z=1}^{Z} p(b|a, \vartheta_z) \quad (14)$$

Where, Z is the number of independent neural networks.

3.2.2 Model Pruning and Quantization

Model pruning declines the model size by increasing the sparsity of the parameters as removing irrelevant parameters. In conventional pruning, the trained model is iteratively pruned until it attains the desired sparsity, and after pruning, the model with pruned parameters is retrained. The considered ResNet model has a function F $(u;\ w)$, then pruned model has a function $F(u; w \odot Z)$, where $Z \in \{0, 1\}$ is a pruning mask to remove the parameter. In practice, the pruned values of w parameters are either nullified or eliminated. Model quantization is a model compression technique that converts the parameter representation from floating-point 32 bits to 8 bits or fewer. Quantization is complementary to pruning techniques and is harmless for accuracy. Thus, the deep ensemble is pruned and quantized to decrease the linearly increasing size.
A simple strategy is used for training from the beginning. The Z shared ResNet structure is randomly initialized and independently train each model using the whole training set until achievable accuracy is attained. Thereafter, the ResNet model is pruned until each model reached the required sparsity by using magnitude-based weight pruning, and a few epochs retrain the pruned networks. At last, each model's weight is quantized and 32-bit floating point representation is converted to an 8-bit representation. This lightweight deep ensemble model is determined by applying pruning and quantization which is robust to unknown instances during the inference time.

The loss function in deep ensembles is evaluated using three predictive scoring rules like NLL, Brier score and compression ratio in the distributed image which is given below.

- Brier score is computed as the squared error of the one-hot encoded true value and the predicted probability.

$$Brier\ score = \frac{1}{m}\frac{1}{i}\sum_{k=1}^{m}\sum_{i=1}^{I}(1[b_k^* = i] - p(b = i|u_k)) \quad (15)$$

- Compression ratio is the ratio of uncompressed size to compressed size

$$Compression\ ratio\ = \frac{uncompressed\ size}{compressed\ size} \quad (16)$$

In addition, the Ensemble disagreement and Ensemble ambiguity is used to measures the disagreement among the ensemble members and also to measure the uncertainty. The Ensemble ambiguity is originally the variance of the weighted predictions of the ensemble around the weighted mean prediction which helps the measure the uncertainty in deep ensembles.

$$Ensemble\ disagreement\ = \sum_{z=1}^{Z}(p_{\vartheta_z}(b|a)||\ p_{\mathcal{E}}\ ((b|a))^2 \quad (17)$$

Where, $p_{\mathcal{E}}\ ((b|a) = Z^{-1}\sum_z^Z p_{\vartheta_z}(b|a)$

$$Ensemble\ disagreement\ = \sum_{z=1}^{Z}(p_{\vartheta_z}(b|a) -\ p_{\mathcal{E}}\ ((b|a))^2 \quad (18)$$

Although, the deep ensemble are uniformly weighted; the ensemble ambiguity in this does not count the ensemble members' weight. Such measures should be increased in the inference image that comes from another distribution from the training image distribution. The schematic representation of this proposed model is depicted in figure 2.

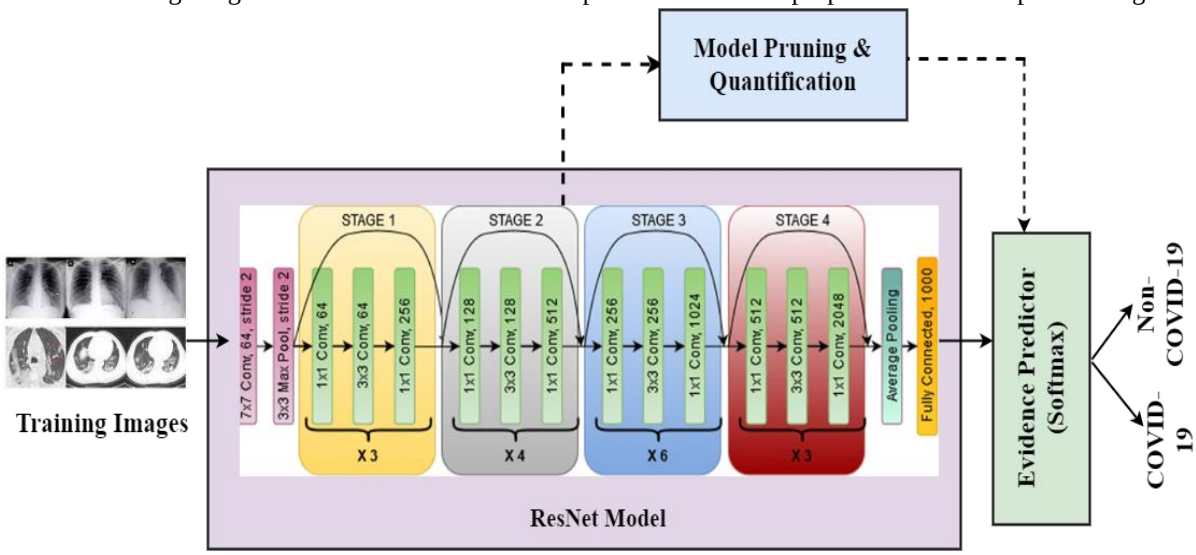


Figure 2 Schematic Representation of the proposed EECDE–ResNet model

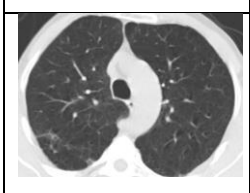
Hence, the proposed EECDE – ResNet model is deployable to memory-intensive and uncertainty-aware tasks for the efficient prediction of COVID-19 detection tasks using CT and CXR images.

4. RESULT AND DISCUSSION

4.1 Dataset Description

For the experimental analysis, the datasets are collected from different websites which are listed below.

CT images: SARS-CoV-2 CT dataset is collected from [34] which consists of 1252 Covid-19 positive images and 1230 Covid-19 negative images, totally 2482 CT images. The Normal and COVID-19 images from CT scans is depicted in Figure 3. For the experimental purpose, 60% (751 images form Covid-19 positive images and 738 images from Covid-19 negative images) are used for training and the remaining 40% (500 images form Covid-19 positive images and 492 images from Covid-19 negative images) are used for testing

Normal	COVID-19	Normal	COVID-19
			

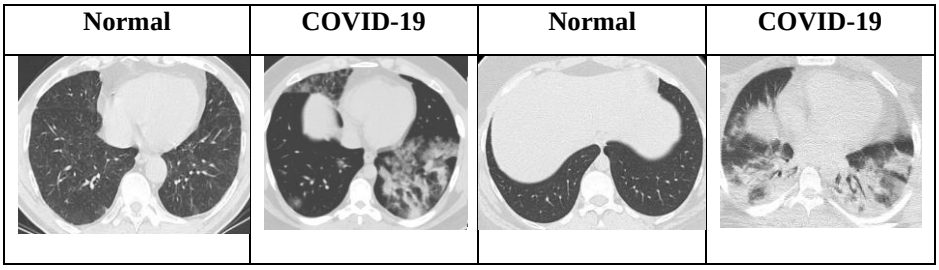


Figure 3 CT images for normal and Covid-19 cases

CXR Images: Database of Covid-19 Radiography is collected from [35]. The first version of dataset consists of 219 COVID-19 and 1341 normal images. The second versions of dataset consists of 3616 COVID-19 and 10,192 normal images. Totally 3835 COVID- 19 images and 11533 normal images available in the dataset. Since, the balanced images for both classes required for training and testing, totally, 3835 COVID-19 and 3835 normal images are considered for experiments. For the training, 60% (2301 images form Covid-19 cases and 2301 from normal cases) of images are used. For testing, 40% (1534 images form Covid-19 cases and 1534 from normal cases) of images are used. Figure 4 depicts the Normal and COVID-19 images from CXR image dataset

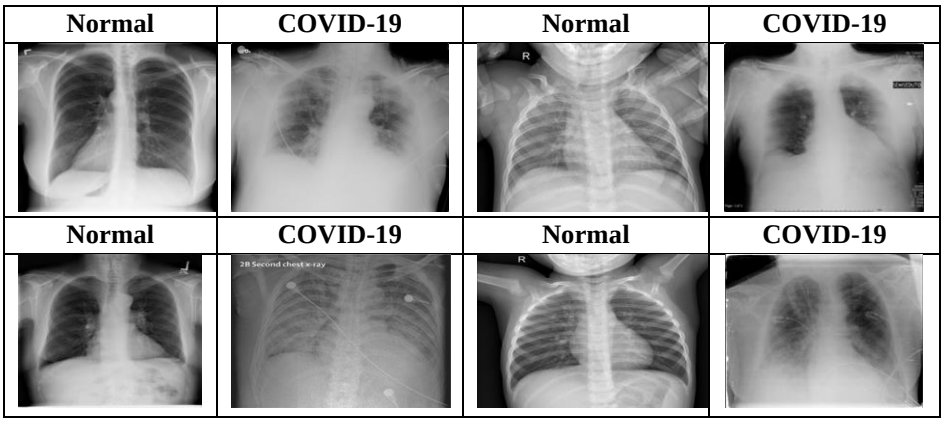


Figure 4 CXR images for normal and Covid-19 cases

4.2 Performance Evaluation

The performance of proposed EECDE – ResNet model and existing classification algorithms like LightEfficientNetV2 [21], SE-CNN [22], DTL-CNN [27], LW-CovidNet [28], DeepCOVNet [29], DDCNN [30] ESMO- ResNet [18] and ECBRM - ResNet [19] are implemented in MATLAB 2017b for detecting SARS-CoV-2 using CT and CXR dataset. The below Table 1 and Table2 provides the confusion matrix for CT images and CXR images respectively. True Positive (TP) and True Negative (TN) solutions are those in which the classifier predicts the covid-19 and normal situations as themselves, respectively. False Positive (FP) solutions occur when the classifier incorrectly identifies a normal cases as Covid-19 instances, whereas False Negative (FN) responses occur when the classifier incorrectly predicts the Covid-19 as normal.

Table 1. Confusion Matrix for CT and CXR images

	CT Images				CXR Images			
Methods	**TP**	**TN**	**FP**	**FN**	**TP**	**TN**	**FP**	**FN**
Light-EfficientNetV2	422	405	82	83	1286	1285	248	249
SE-CNN	428	413	75	76	1303	1302	231	232
DTL-CNN	437	422	67	66	1314	1313	220	221
LW-CovidNet	443	426	62	61	1326	1325	209	210
DDCNN	448	433	56	55	1349	1348	185	186
DeepCOVNet	453	442	49	48	1366	1365	168	169

ESMO-ResNet	457	448	43	44	1382	1381	152	153
ECRBM-ResNet	472	471	21	28	1444	1443	88	92
EECDE – ResNet	488	487	9	8	1490	1489	45	44

4.2.1 Accuracy

Accuracy is an essential parameter for assessing classification systems, and it is characterized as the ratio of accurate COVID-19 forecasts to entire observations conducted by using the provided data sample. The accuracy percentage can be estimated as in Eq. (19),

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \quad (19)$$

Figure 5 demonstrates the accuracy values obtained by Light-EfficientNetV2, SE-CNN, DTL-CNN, LW-CovidNet, DDCNN, DeepCOVNet, ESMO-ResNet, ECRBM-ResNet and EECDE – ResNet for Covid-19 identification using SARS-CoV-2 CT scan dataset and Covid-19 Radiography dataset. This observation shows that the proposed EECDE – ResNet model outperforms the other classical models for the better COVID-19 identification. For instances, the accuracy of EECDE – ResNet is 17.92% and 16.71% greater than Light-EfficientNetV2, 15.93% and 15.05% greater than SE-CNN, 13.51% and 14.08% greater than DTL-CNN, 12.20% and 13.13%, greater than LW-CovidNet, 10.67% and 11.12% greater than DDCNN, 8.94% and 9.73% greater than DeepCOVNet, 7.73% and 8.39% greater than ESMO-ResNet, 3.39% and 3.78%, greater than ECRBM-ResNet for CT and CXR images, respectively.

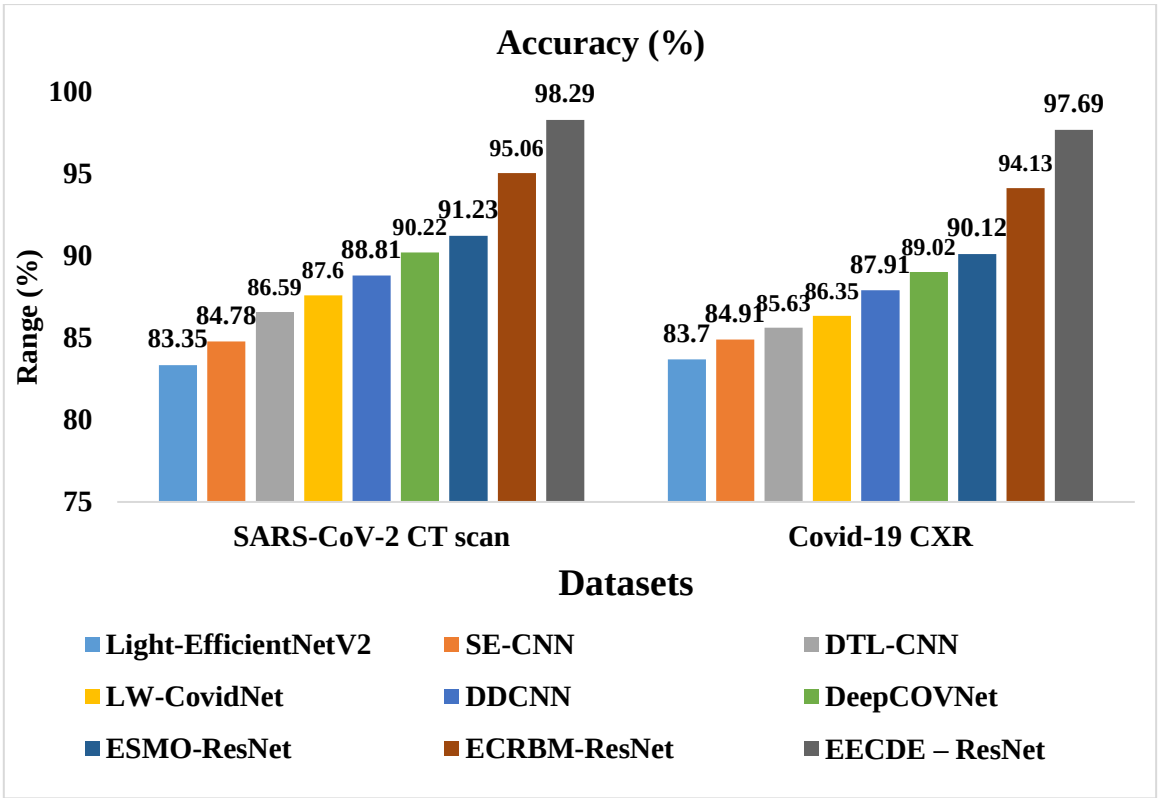


Figure 5 Accuracy Comparison for proposed and existing methods

4.2.2 Precision

It is the ratio of exactly classified categories of Covid-19 cases at TP and FP rates.

$$Precision = \frac{TP}{TP + FP} \quad (20)$$

Figure 6 shows the precision values obtained for Covid-19 identification using the SARS-CoV-2 CT scan dataset and Covid-19 Radiography dataset using Light-EfficientNetV2, SE-CNN, DTL-CNN, LW-CovidNet, DDCNN, DeepCOVNet, ESMO-ResNet, ECRBM-ResNet, and EECDE - ResNet models. This investigation demonstrates that for better COVID-19 identification, the suggested EECDE - ResNet model performs better than the other traditional methods. For instances, the accuracy of EECDE-ResNet is 17.26% and 16.75% higher than Light-EfficientNetV2, 15.39% and 14.96% higher than SE-CNN, 13.23% and 13.99% higher than DTL-CNN, 11.93% and 13.04% higher than LW-CovidNet, 10.46% and 11.04% higher than DDCNN, 8.80% and 9.65% higher than DeepCOVNet, 7.42% and 8.39% higher than ESMO-ResNet, 2.55% and 3.59% higher than ECRBM-ResNet in the case of CT and CXR images, respectively.

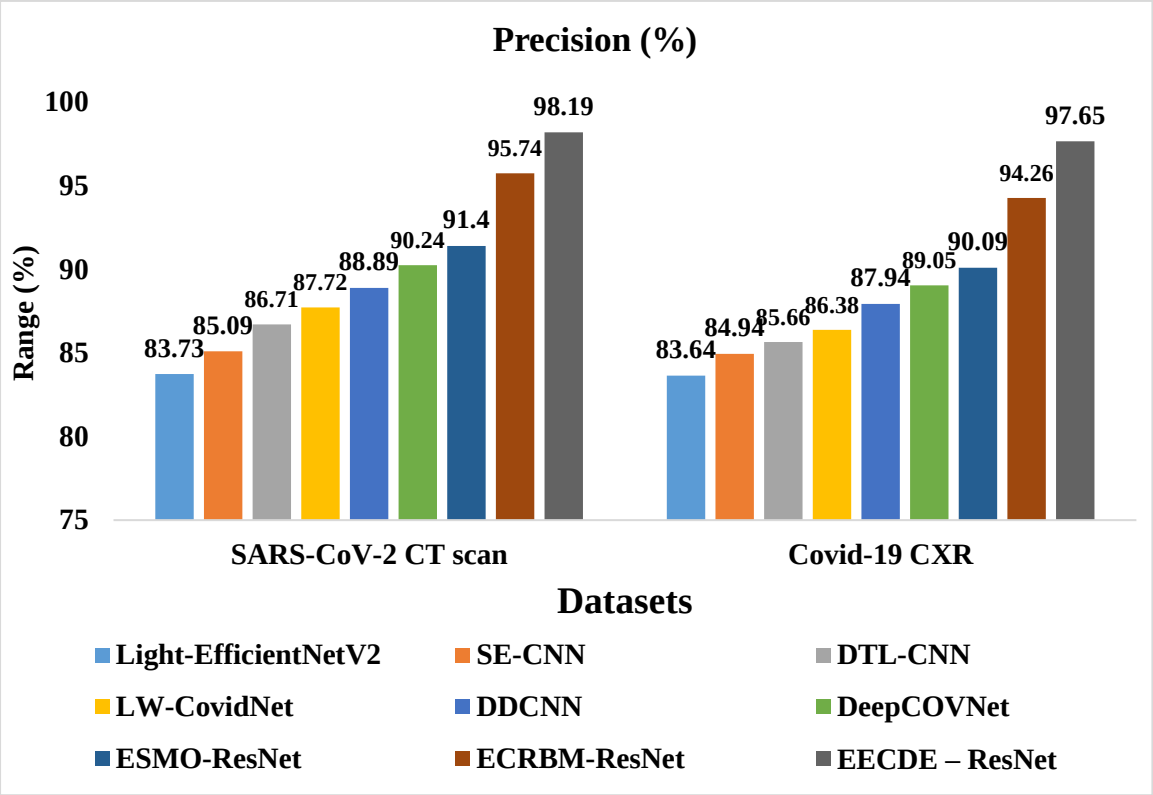


Figure 6 Precision Comparison for proposed and existing methods

4.2.3 Recall

It is the ratio of exactly classified categories of Covid-19 cases at TP and FN rates.

$$Recall = \frac{TP}{TP + FN} \quad (21)$$

Figure 7 depicts the recall values achieved by Light-EfficientNetV2, SE-CNN, DTL-CNN, LW-CovidNet, DDCNN, DeepCOVNet, ESMO-ResNet, ECRBM-ResNet, and EECDE - ResNet for Covid-19 identification using the SARS-CoV-2 CT scan dataset and the Covid-19 Radiography dataset for Light-EfficientNetV2. According to this investigation, the suggested EECDE - ResNet model exceeds than other classical models for improved COVID-19 identification. For instances, the accuracy of EECDE - ResNet is 17.74% and 16.90%, greater than Light-EfficientNetV2, 15.86% and 15.11%, greater than SE-CNN, 13.24% and 14.15%, greater than DTL-CNN, 11.93% and 13.19%, greater than LW-CovidNet, 10.46% and 11.19%, greater than DDCNN, 8.81% and 9.81% greater than DeepCOVNet, 7.86% and 8.39%, greater than ESMO-ResNet, 4.22% and 3.94%, greater than ECRBM-ResNet for CT and CXR images, respectively.

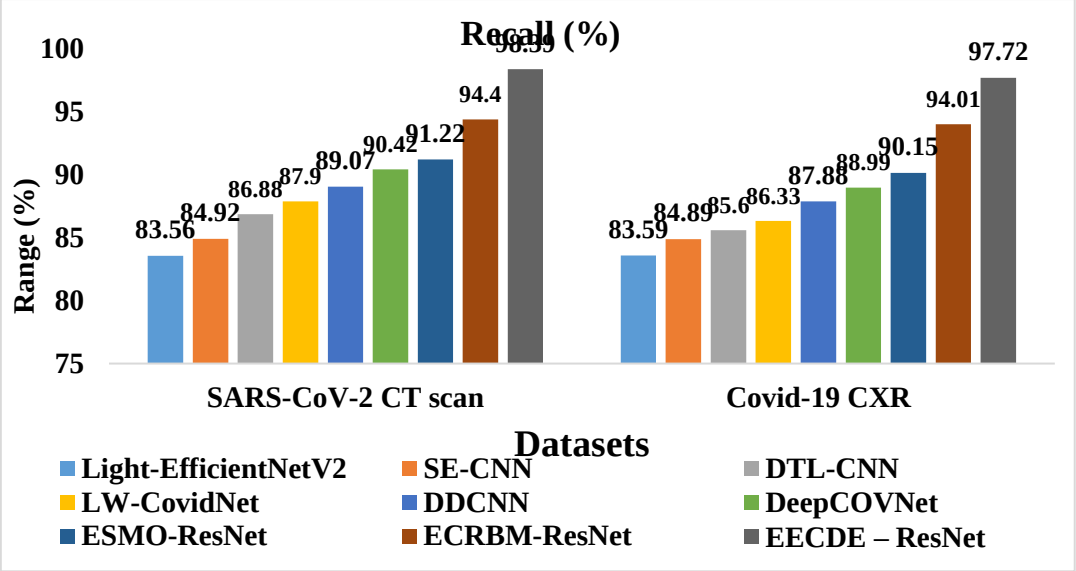


Figure 7 Recall Comparison for proposed and existing methods

4.2.4 F-Measure

It is defined as the weighted average of precision and recall, with a value of '1' being the highest and a value of '0' representing the lowest.

$$F - measure = 2 \times \frac{Precision \cdot Recall}{Precision + Recall} \quad (22)$$

Using the SARS-CoV-2 CT scan dataset and the Covid-19 Radiography dataset, the F-Measure values produced by the Light-EfficientNetV2, SE-CNN, DTL-CNN, LW-CovidNet, DDCNN, DeepCOVNet, ESMO-ResNet, ECRBM-ResNet, and EECDE-ResNet for Covid-19 identification are shown in figure 9. According to the results of this observation form figure 9, the proposed EECDE - ResNet model provides superior COVID-19 identification compared to the other traditional models considered. For example, the accuracy of EECDE - ResNet is 17.50% and 16.84% higher than Light-EfficientNetV2, 15.63% and 15.05% higher than SE-CNN, 13.25% and 14.08% higher than DTL-CNN, 11.93% and 13.11% higher than LW-CovidNet, 10.46% and 11.12% highe than DDCNN, 8.81% and 9.73% higher than DeepCOVNet, 7.64% and 8.39% higher than ESMO-ResNet, 3.38% and 3.78% higher than ECRBM-ResNet for CT and CXR images respectively.

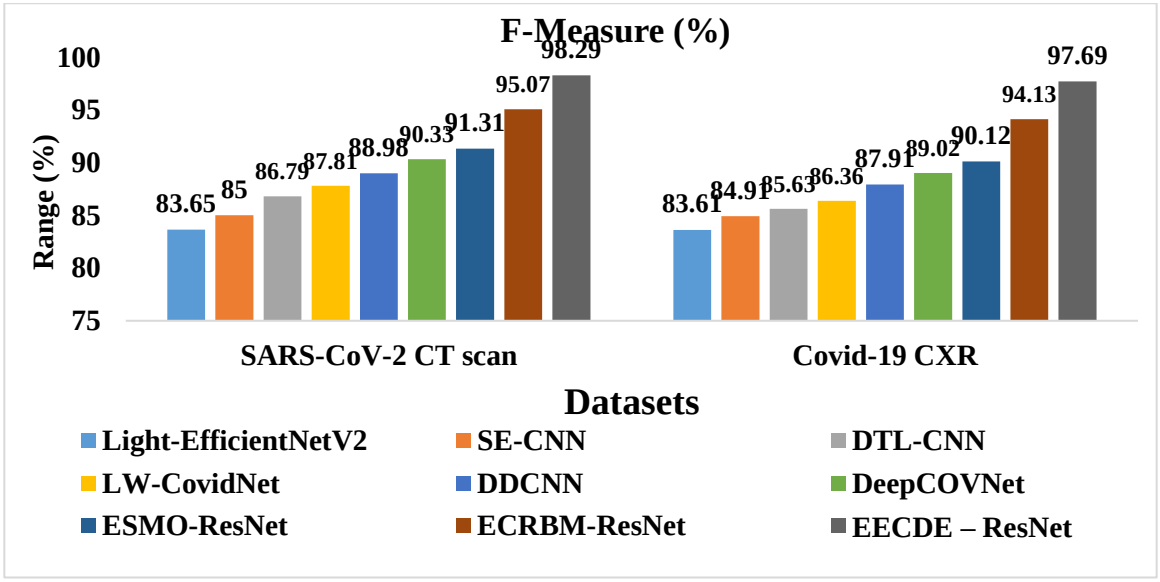


Figure 8 F1-Measure Comparison for proposed and existing methods

The above performance comparison proves that the proposed EECDE – ResNet model results in higher results while compared to other existing models which resolves the problems of uncertainty estimation and memory intensive on provided training images (CT and X-ray) for COVID-19 detection

5. CONCLUSION

In this paper, the EECDE – ResNet model is developed to eliminate the uncertainty issues and memory intensive workloads in the constructed classifier to provide accurate results in COVID-19 detection tasks using CT and CXR images. In this model, the theory of evidence perceptive is used to resolve the uncertainty issue. The softmax layers of ResNet models is altered with Dirichlet density parameters to represent learner predictions as a distribution over possible softmax outputs (Evidence predictor). To reduce the challenges in memory-intensive tasks, model pruning and quantization techniques are performed along with deep ensemble and analyzed the effect in the context of uncertainty metrics. Increased disagreement in deep ensemble models implies increased uncertainty which helps in making more robust predictions for COVID-19. At last, the investigational outcomes show that the EECDE – ResNet model on the CT and CXR images dataset attains 98.29% and 97.69% accuracy compared to the cutting-edge models for COVID-19 detection and diagnosis.

REFERENCES

[1] Huang, C., Wang, Y., Li, X., Ren, L., Zhao, J., Hu, Y., ... & Cao, B. (2020). Clinical features of patients infected with 2019 novel coronavirus in Wuhan, China. The lancet, 395(10223), 497-506.

[2] Lu, H., Stratton, C. W., & Tang, Y. W. (2020). Outbreak of pneumonia of unknown etiology in Wuhan, China: The mystery and the miracle. Journal of medical virology, 92(4), 401.

[3] Chen, N., Zhou, M., Dong, X., Qu, J., Gong, F., Han, Y., ... & Zhang, L. (2020). Epidemiological and clinical characteristics of 99 cases of 2019 novel coronavirus pneumonia in Wuhan, China: a descriptive study. The lancet, 395(10223), 507-513.

[4] Lefkowitz, E. J., Dempsey, D. M., Hendrickson, R. C., Orton, R. J., Siddell, S. G., & Smith, D. B. (2018). Virus taxonomy: the database of the International Committee on Taxonomy of Viruses (ICTV). Nucleic acids research, 46(D1), D708-D717.

[5] Girdhar, A., Kapur, H., Kumar, V., Kaur, M., Singh, D., & Damasevicius, R. (2021). Effect of COVID-19 outbreak on urban health and environment. Air Quality, Atmosphere & Health, 14, 389-397.

[6] Corman, V. M., Landt, O., Kaiser, M., Molenkamp, R., Meijer, A., Chu, D. K., ... & Drosten, C. (2020). Detection of 2019 novel coronavirus (2019-nCoV) by real-time RT-PCR. Eurosurveillance, 25(3), 2000045.

[7] Chaimayo, C., Kaewnaphan, B., Tanlieng, N., Athipanyasilp, N., Sirijatuphat, R., Chayakulkeeree, M., ... & Horthongkham, N. (2020). Rapid SARS-CoV-2 antigen detection assay in comparison with real-time RT-PCR assay for laboratory diagnosis of COVID-19 in Thailand. Virology journal, 17, 1-7.

[8] Varadarajan, V., Shabani, M., Ambale Venkatesh, B., & Lima, J. A. (2021). Role of imaging in diagnosis and management of COVID-19: a multiorgan multimodality imaging review. Frontiers in Medicine, 2013.

[9] Rahmani, A. M., Azhir, E., Naserbakht, M., Mohammadi, M., Aldalwie, A. H. M., Majeed, M. K., ... & Hosseinzadeh, M. (2022). Automatic COVID-19 detection mechanisms and approaches from medical images: a systematic review. Multimedia tools and applications, 81(20), 28779-28798.

[10] Soomro, T. A., Zheng, L., Afifi, A. J., Ali, A., Yin, M., & Gao, J. (2022). Artificial intelligence (AI) for medical imaging to combat coronavirus disease (COVID-19): A detailed review with direction for future research. Artificial Intelligence Review, 1-31.

[11] Chatterjee, P., Biswas, M., & Das, A. K. (2021, February). Specialized covid-19 detection techniques with machine learning. In Journal of Physics: Conference Series (Vol. 1797, No. 1, p. 012033). IOP Publishing.

[12] Guhathakurata, S., Kundu, S., Chakraborty, A., & Banerjee, J. S. (2021). A novel approach to predict COVID-19 using support vector machine. In Data Science for COVID-19 (pp. 351-364). Academic Press.

[13] Prasannavenkatesan, T., Jacob, I. J., Ruby, A. U., & Vamsidhar, Y. (2021). Prediction of COVID-19 Possibilities using KNN Classification Algorithm. Accident Analysis and Prevention. Int J Cur Res Rev, 13(06), 156.

[14] Shanbehzadeh, M., Nopour, R., & Kazemi-Arpanahi, H. (2022). Design of an artificial neural network to predict mortality among COVID-19 patients. Informatics in medicine unlocked, 31, 100983.

[15] Subramanian, N., Elharrouss, O., Al-Maadeed, S., & Chowdhury, M. (2022). A review of deep learning-based detection methods for COVID-19. Computers in Biology and Medicine, 105233.
[16] Khan, W. Z., Azam, F., & Khan, M. K. (2022). Deep Learning Based COVID-19 Detection: Challenges and Future Directions. IEEE Transactions on Artificial Intelligence.
[17] Shamsi A, Asgharnezhad H, Jokandan SS, Khosravi A, Kebria PM, Nahavandi D, Nahavandi S, Srinivasan D. An Uncertainty-Aware Transfer Learning-Based Framework for COVID-19 Diagnosis. IEEE Trans Neural Netw Learn Syst. 2021 Apr;32(4):1408-1417. doi: 10.1109/TNNLS.2021.3054306. Epub 2021 Apr 2. PMID: 33571095.
[18] Deepika, S., and Rajeshwari, M. (2023). An optimized uncertainty aware fine-tuned transfer learning for COVID-19 diagnosis from medical images
[19] Deepika, S., and Rajeshwari, M. (2023). ResNet Pertained Model and Enhanced Convolution Restricted Boltzmann Machine by Contrastive Divergence for Lung Disease Prediction.
[20] Biradar, V. G., Alqahtani, M. A., Nagaraj, H. C., Ahmed, E. A., Tripathi, V., Botto-Tobar, M., & Atiglah, H. K. (2022). An Effective Deep Learning Model for Health Monitoring and Detection of COVID-19 Infected Patients: An End-to-End Solution. Computational Intelligence and Neuroscience, 2022.
[21] Chouat, I., Echtioui, A., Khemakhem, R., Zouch, W., Ghorbel, M., & Hamida, A. B. (2022). COVID-19 detection in CT and CXR images using deep learning models. Biogerontology, 23(1), 65-84.
[22] Deb, S. D., Jha, R. K., Jha, K., & Tripathi, P. S. (2022). A multi model ensemble based deep convolution neural network structure for detection of COVID19. Biomedical Signal Processing and Control, 71, 103126.
[23] Huang, M. L., & Liao, Y. C. (2022). A lightweight CNN-based network on COVID-19 detection using X-ray and CT images. Computers in Biology and Medicine, 146, 105604.
[24] Kogilavani, S. V., Prabhu, J., Sandhiya, R., Kumar, M. S., Subramaniam, U., Karthick, A., ... & Imam, S. B. S. (2022). COVID-19 detection based on lung CT scan using deep learning techniques. Computational and Mathematical Methods in Medicine, 2022.
[25] Zouch, W., Sagga, D., Echtioui, A., Khemakhem, R., Ghorbel, M., Mhiri, C., & Hamida, A. B. (2022). Detection of COVID-19 from CT and chest X-ray images using deep learning models. Annals of Biomedical Engineering, 50(7), 825-835.
[26] Gour, M., & Jain, S. (2022). Automated COVID-19 detection from X-ray and CT images with stacked ensemble convolutional neural network. Biocybernetics and Biomedical Engineering, 42(1), 27-41.
[27] Kathamuthu, N. D., Subramaniam, S., Le, Q. H., Muthusamy, S., Panchal, H., Sundararajan, S. C. M., ... & Zahra, M. M. A. (2023). A deep transfer learning-based convolution neural network model for COVID-19 detection using computed tomography scan images for medical applications. *Advances in Engineering Software*, *175*, 103317.
[28] Ahmed, N., Tan, X., & Ma, L. (2023). LW-CovidNet: Automatic covid-19 lung infection detection from chest X-ray images. IET Image Processing, 17(2), 362-374.
[29] Bhattacharjee, V., Priya, A., Kumari, N., & Anwar, S. (2023). DeepCOVNet Model for COVID-19 Detection Using Chest X-Ray Images. Wireless Personal Communications, 1-18.
[30] Foysal, M., Hossain, A. B. M., Yassine, A., & Hossain, M. S. (2023). Detection of COVID-19 Case from Chest CT Images Using Deformable Deep Convolutional Neural Network. Journal of Healthcare Engineering, 2023.
[31] Dempster, A. P. (2008). Upper and lower probabilities induced by a multivalued mapping. Classic works of the Dempster-Shafer theory of belief functions, 219(2), 57-72.
[32] Jøsang, A. (2016). Subjective Logic: A Formalism for Reasoning Under Uncertainty. Springer.
[33] Cheng, Y., Wang, D., Zhou, P., & Zhang, T. (2017). A survey of model compression and acceleration for deep neural networks. arXiv 2017. arXiv preprint arXiv:1710.09282.
[34] https://www.kaggle.com/plameneduardo/sarscov2-ctscan-database.
[35] https://www.kaggle.com/datasets/tawsifurrahman/covid19-radiography-database.