EPiC Series in Computing

Volume 101, 2024, Pages 58–72

Proceedings of the 16th International Conference on Bioinformatics and Computational Biology (BICOB-2024)

EPiC Computing

Generalized Deep Learning Models for COVID-19 Detection with Transfer and Continual Learning

Richard Annan[1], Hong Qin[2], and Letu Qingge[1*]

[1] Department of Computer Science, North Carolina A&T State University, Greensboro, North Carolina, USA
rkannan@aggies.ncat.edu, lqingge@ncat.edu

[2] Deparment of Computer Science and Engineering, University of Tennessee at Chattanooga, Chattanooga, Tennessee, U.S.A.
hong-qin@utc.edu

Abstract

Deep learning has achieved great success for detecting COVID-19 from CT scan images. However, there is lack of generalization ability for the existing models. For example, one model with a higher prediction accuracy developed on one dataset cannot be used to predict on another dataset. Thus, developing a robust deep learning model that has a great generalization ability is a significant need. In this paper, we first apply three deep learning models, namely convolutional neural network (CNN), capsule neural network (CapsNet) and vision transformer (ViT) and test their generalization abilities. Then, we develop and hypertune the models based on transfer learning to generalize the model performance on new datasets. However, the transfer learning technique always has the catastrophic forgetting issue which lead to lower prediction accuracy on its original training dataset. Lastly, we will apply continual learning based on modified elastic weight consolidation (EWC) regularization technique to address the catastrophic forgetting issue and improve the models' prediction accuracy on both new and original training datasets. Our results on cross-data validation show that our proposed models not only achieve better prediction accuracy of up to 97.85% compared with the existing state-of-the-art models, but also the proposed models with EWC show great generalization ability and retain the higher prediction accuracy on both new dataset and the training dataset. Extensive experiments show that our proposed COVID-CNN model with EWC outperforms ViT and CapsNet with an impressive 82.26% knowledge retention rate on the original training dataset. Our developed code can be found from https://github.com/astonish24/-QinggeLab_BICOB24.

Keywords: COVID-19 Detection; Deep Learning; Transfer Learning; Continual Learning

*This work is supported by the U.S. National Science Foundation under Award 2200138. (Corresponding author: Letu Qingge. Email: lqingge@ncat.edu)

1 Introduction

The COVID-19 pandemic has significantly impact global health with over 469.9 million cases and 6.1 million deaths reported by the WHO as of March 20, 2022 [14]. It has strained healthcare systems, resulting in overburdening medical resources and increasing mortality rates. This situation highlights the urgent need for effective treatments and vaccines. The Reverse Transcription Polymerase Chain Reaction (RT-PCR) test, widely used for COVID-19 detection, has faced criticism for its inconsistent sensitivity, leading to concerns about its reliability as a diagnostic tool for COVID-19 [10]. False-negative results, influenced by various factors, would bring negative impacts for public health and healthcare delivery. This underscores the necessity for alternative diagnostic methods to improve the accuracy of COVID-19 detection. Also, due to the limitations of RT-PCR testing for COVID-19 detection, the chest Computed Tomography (CT) scans have emerged as a viable option. CT scan images play a crucial role in identifying distinctive COVID-19 pneumonia patterns [14].

Deep learning models including Convolutional Neural Networks (CNNs), Capsule Neural Networks (CapsNet) and Vision Transformers (ViT) provide effective solutions for improving the accuracy of COVID 19 detection from CT scans by distinguishing it from diverse respiratory diseases. These models utilize feature extraction, attention mechanism and transfer learning to enhance precision and monitor disease progression across healthcare data [10, 21]. However, most of the existing deep learning models are lack of generalization ability. These models become excessively specialized on particular datasets, leading to challenges in performing effectively on new unseen data. This challenge is further complicated by the sources of COVID-19 datasets and the limited availability of labeled training data for rare disease like COVID-19. Therefore, it is essential to develop robust deep learning models that adapt to variations of the virus to ensure accurate and reliable COVID-19 detection across different virus strains and data sources.

In this paper, we focus on developing robust deep learning models with a great generalization ability by utilizing transfer and continual learning techniques on diverse CT scan datasets derived from various healthcare settings. The approach includes using CNN, CapsNet and ViT models for COVID-19 detection from CT scan images and then employing transfer learning to generalize the model prediction ability on new dataset. One major challenge in this process is catastrophic forgetting, where the model performance decline on the original training dataset. To address this issue, a continual learning strategy is implemented based on modified elastic weight consolidation (EWC) technique. This allows the models to retain knowledge from old dataset and effectively adapt to the new dataset. This paper is organized as follows. In section 2, we introduce the relevant works, offering a foundational understanding and gaps in the existing works regarding the utilization of deep learning for COVID-19 detection from CT scan images. In section 3, we explain our main contributions with the design of the models, the formulation of catastrophic forgetting problem, and the application of transfer learning combined with continual learning to tackle the catastrophic forgetting issue. Section 4 lays out our research outcomes and analysis results on different COVID-19 datasets. We conclude the paper in the last section.

2 Related Works

Computed tomography (CT) scans have emerged as an important tool in the assessment of individuals affected by COVID-19. Artificial Intelligence (AI) and deep learning models exhibit remarkable proficiency in the analysis of medical images including CT scans to uncover latent patterns and attributes. In this section, we introduce AI and deep learning applications for COVID-19 detection using CT scan images. Identifying the strengths and limitations of the existing works will lead us to address the need of building robust models with a great generalization ability for COVID-19 detection from CT scan images.

Numerous machine learning models have been developed for detecting COVID-19 using CT scan images, such as support vector machine, Bayesian models, and decision tree [9]. These methods with the manual feature extraction and dimension reduction techniques like principal component analysis (PCA) achieved better results [24]. Deep learning models, particularly CNNs, have shown promising results in analyzing medical images. Since medical images are vulnerable to noises, which are random variations in pixel intensity or voxel values that distort the true information present in the image, anisotropic diffusion was shown effective in reducing noises in medical images when it was applied before feeding them in CNN models [4, 23]. However, high levels of contamination still can affect CNN model performance. A novel filter called layer discrimination with max/min intensities elimination was proposed for detecting and removing Impulse/Poisson noises in CT images [11], which improves image clarity and reduces blurriness, therefore improving the performance of assessment fusion based CNN models. Thus, the preprocessing of CT scan images is crucial in enhancing the performance of CNN models.

However, CNN models also have limitations in analyzing CT scans, particularly in detecting sub-visual lesions, which impacts their effectiveness in medical decision support. To address this, researchers designed Deep-LungParenchyma-Enhancing (DLPE) [27], a novel deep learning workflow to identify subtle abnormalities in CT scans of COVID-19 patients to improve model performance and minimize false discoveries. Xiao *et al.* [24] introduced PAM-DenseNet, a convolutional neural network (CNN) specifically designed to handle coarse labels without the need for manual delineation of infection regions by incorporating a parallel attention module that effectively emphasizes crucial features, leading to improved focus on infection regions. Miao *et al.* [16] analyzed receptive fields at the individual layer in CNN model and revealed certain layers' contributions to overall model accuracy. This observation highlights a fundamental challenge inherent in CNNs. Their study confirmed the correlation between the size of the image and the size of the receptive field significantly impacts the model performance. Richter et al. [18] showed the fixed receptive fields may insufficiently capture the complex interrelationships and broader context within medical images. This limitation has the potential to result in suboptimal predictive performance, particularly in the analysis of COVID-19 CT scan images.

Furthermore, emerging deep learning paradigms, namely Capsule Networks (CapsNets) and Vision Transformers (ViTs), provided promising alternatives for COVID-19 CT scan images analysis. CapsNets pioneered by Geoffrey Hinton, play a significant advancement in neural network architecture [19]. Unlike traditional CNNs, CapsNets employ dynamic routing to efficiently route lower-level features to corresponding higher-level capsules. A capsule is defined as a mathematical vector detecting hierarchical relationships between parts of objects in an image, including its existence, spatial location, dimensions, and orientation. In CapsNets, this architectural innovation replaces conventional neurons with a unique arrangement of capsules, offering a novel approach to feature extraction and routing. Notably, CapsNets have demonstrated effectiveness in the analysis of medical images, particularly for COVID-19 CT scan images [1]. In a similar vein, Vision Transformer (ViT) was also successfully designed for computer vision tasks, such as image classification and object detection. Different from traditional CNNs, ViTs rely on transformer architectures that were originally developed for natural language processing tasks [7]. Transformers use self-attention mechanisms to capture relationships between different elements in a sequence. CapsNet and ViT show potential in addressing the limitations of CNNs, improving the accuracy and interpretability of predictions in the analysis of COVID-19 CT scan images. However, deep learning models often encounter challenges in effectively generalizing to data that not included in their training process. Such challenges result in a decline in terms of model performance when evaluated on distinct datasets, which is not part of the initial training data [5], especially for COVID-19 CT scans images. Thus, it is crucial to develop robust models with a great generalization ability beyond the specific datasets used for training. Next, we will discuss our proposed deep learning models and techniques to build robust model with generalization ability for COVID-19 detection.

3 Methodology

Our main goal in this paper is to build robust deep learning models with great generalization ability for COVID-19 detection. In this section, we employ CNN, CapsNet and ViT models for COVID-19 CT scan images analysis. To ensure our proposed models have better generalization ability, we use transfer learning technique to test model performance on the new dataset. However, we observe that transfer learning always suffer from catastrophic forgetting issues, which is a substantial decline of model performance on its previous training dataset. To solve this, we use continual learning based on modified elastic weight consolidation (EWC) regularization technique to improve model prediction accuracy. To the best of our knowledge, our paper is the first one using transfer learning with EWC technique to explore model generalization ability for COVID-19 detection from CT scan images.

3.1 Datasets and Preprocessing

The training dataset we will use in this paper is collected from the radiologist centers of teaching hospitals in Sao Palo (Brazil) [22] and Tehran (Iran) [8]. We combine these two datasets to obtain a balanced dataset and name it as *dataset_mod_dev* for model development. The other dataset, namely *dataset_mod_gen* is collected from countries including Russia, China, Italy, Turkey, and Iran [15] to test model generalization ability. The *dataset_mod_dev* is comprised of 4649 images, of which 2476 are from patients with COVID-19 and 2173 are from patients without COVID-19. The *dataset_mod_gen* is a more diverse dataset used in [15], comprising of 14486 images with 7593 COVID-19 cases and 6893 Non COVID-19 cases. Figure 1 illustrates COVID-19 CT scan images as examples, showing patients with and without COVID-19 infection.

In a preprocessing stage, we apply data augmentation techniques, including rotation, flipping, and zoom adjustments to increase the size of datasets. Images are uniformly resized to dimensions of $300\times300\times1$ pixels, and pixel values are standardized to the interval number of (0, 1), ensuring consistent input data for our proposed models. Additionally, all images within the dataset are shuffled to ensure that the models have a balanced representation of both COVID-19 and non–COVID-19 cases during training. This approach will prevent bias towards any specific class and contribute to model robustness. The *dataset_mod_dev* is divided into training, validation, and test sets, following a ratio of 64:16:20. The *dataset_mod_gen* is split into the ratio of 50:50 for training and testing. We use the validation set for effective model training and hyperparameter tuning, and use the test sets for model evaluation.

3.2 Proposed Models

In this section, we present three models with the modified elastic weight consolidation (EWC) regularization technique to build robust models for COVID-19 detection, aiming to effectively

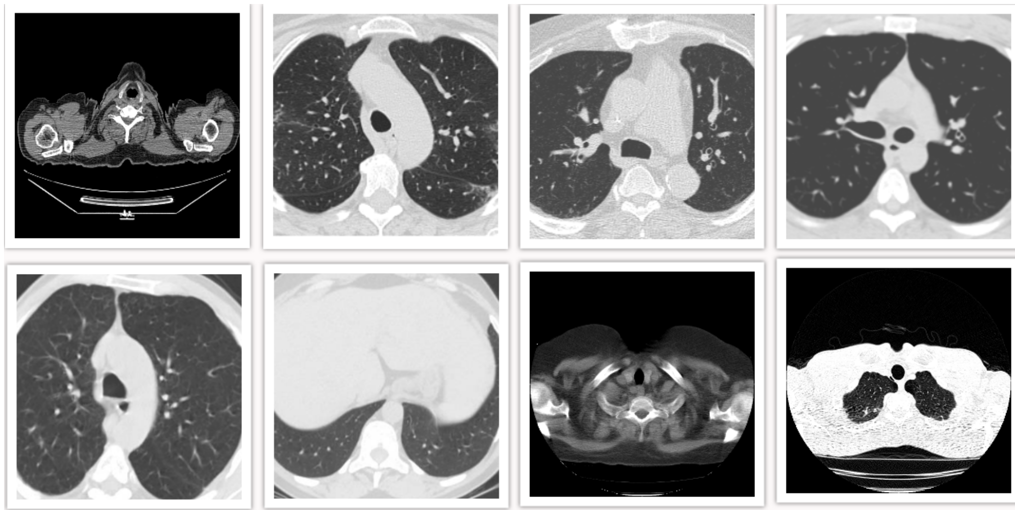


Figure 1: Example Images of Patients with the COVID-19 in the Top Row and Non COVID-19 Cases in the Bottom Row.

balance the acquisition of new knowledge and the preservation of existing information. We build CNN model, referred to as COVID-CNN and apply ViT [19] and CapsNet [3, 20] for COVID-19 detection on CT scan images. The individual model performance is evaluated on both *dataset_mod_dev* and *dataset_mod_gen* datasets. As shown in the Table 1, all proposed models demonstrate higher prediction accuracy on *dataset_mod_dev*, while provide lower prediction accuracy on *dataset_mod_gen*. Next, we will discuss how each model is built for COVID-19 CT scan image analysis.

3.2.1 COVID-CNN Model

The proposed COVID-CNN model is specifically tailored for grayscale images with a dimension of 300×300×1. The initial processing begins with 116 feature maps (filters) used in the first convolutional layer with a kernel size and a stride of 8×8 and 2×2 respectively. Then the output dimension from the first convolutional operation is in the shape of 97×97×116.

After the convolution operation, a stride of 2×2 pooling layer is applied to downsample the feature maps, reducing the spatial dimensions while keeping important information from the previous layer. These operations are crucial for capturing localized and hierarchical patterns of images. Additionally, batch normalization is used for the first convolutional layer to stabilize the training process. The second convolutional layer with batch normalization and max-pooling, also has the same parameters (i.e., 116 filters, a size of 8×8 kernel, and a size of 2×2) as the first one, yielding 116 feature maps with the convolution operation of 10×10×116 shape filters. The next layer output has dimensions of 10×10×116, which serves as the input to the fully connected layer. The fully connected layer comprises of four layers with 362, 184, 78, and 12 neurons, respectively. The ReLU activation function is utilized in these layers, and dropout regularization techniques are applied to address overfitting. The final layer in COVID-CNN is the output layer with a softmax activation function with a size of 2, which ultimately used to classify the output image. The optimal hyperparameter tuning includes a learning rate of 0.001, the ADAM optimizer, and categorical cross-entropy as the loss function. Figure 2 shows our designed COVID-CNN architecture.

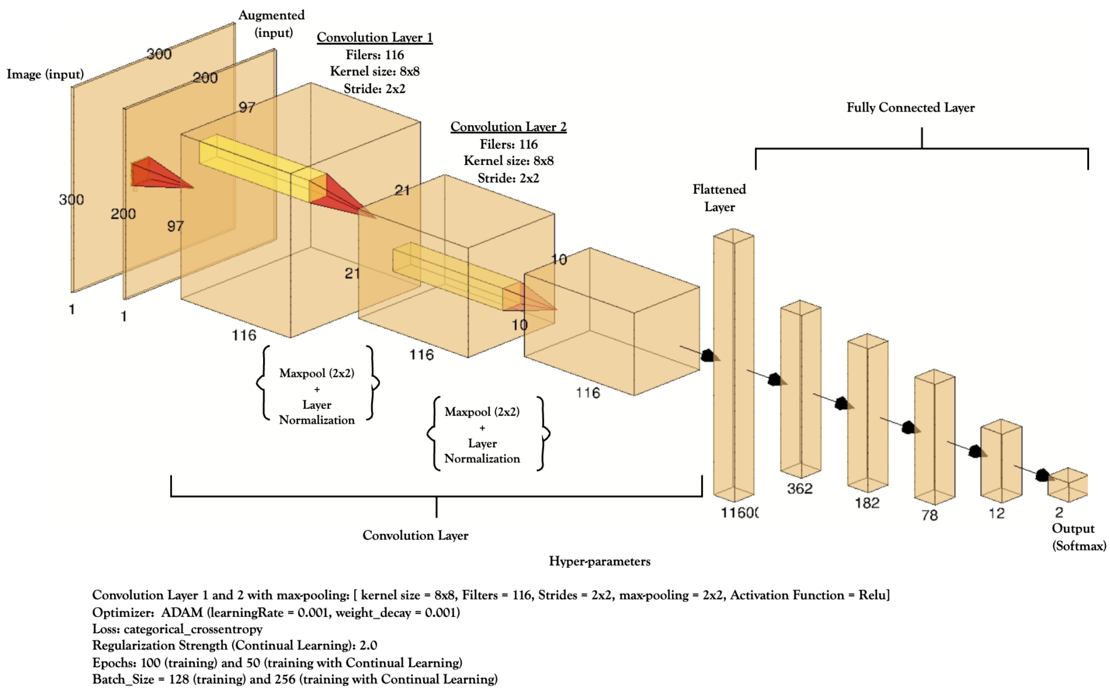
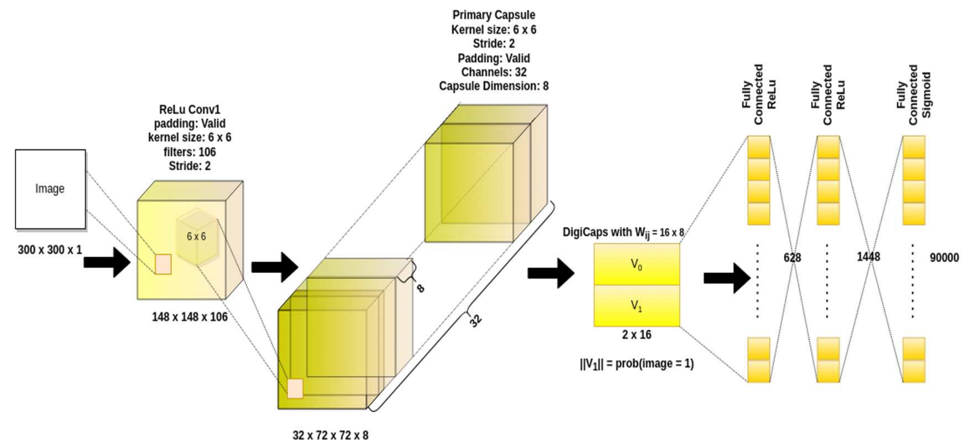


Figure 2: COVID-CNN Architecture

3.2.2 Capsule Network (CapsNet)

The architecture of CapsNet for COVID-19 detection is illustrated in the Figure 3. The input image with dimensions of 300×300×1 is fed through the Conv1 layer with ReLu activation function, which functions as a standard convolutional layer. Within this layer, 106 kernels of size 6×6 and a stride of 2 and no padding are applied, resulting in an output with 106 feature maps. The spatial dimensions are reduced to 148×148, calculated as $\lfloor(300-6)/2\rfloor+1 = 148$. Then, the output is passed into the *PrimaryCapsules* layer, which serves as a modified convolutional layer that supports capsules. Unlike conventional layers that produce scalar outputs, *PrimaryCapsules* generates 8-dimensional vectors. To achieve this, it employs 8×32 kernels to create 32 of 8-dimensional capsules, where 8 output neurons are grouped together to form a capsule. *PrimaryCapsules* utilizes 6×6 kernels with a stride of 2 and no padding to effectively reduce the spatial dimension from 148×148 to 72×72, calculated as $\lfloor(148-6)/2\rfloor+1 = 72$. As a result, within the *PrimaryCapsules*, we have a total of 32×72×72, 8-dimensional capsules. The output from the *PrimaryCapsules* layer is fed into the *DigiCaps* layer, where it undergoes a transformation using a matrix W_{ij} with dimensions of 16×8. This transformation converts the initial 8-dimensional capsule into a 16-dimensional capsule for each class 1 and 2 respectively.

Each *PrimaryCapsules* tries to predict the output of every digit capsule by multiplying its own output vector (u_j) by a weight matrix (W_{ij}). This gives us the prediction vectors $\hat{u}_j|i$ denoted as V_0 or V_1 in Figure 3. The total input s_j to a *DigiCaps* (e.g., V_0 or V_1) is formed



Hyper-parameters

Convolution Layer 1: [kernel size = 6x6, Filters = 106, Strides = 2x2, padding = valid, Activation Function = Relu]
Primary Capsule: [Capsule Dimension = 8, channels = 32, kernel size = 6x6, padding = valid]
DigiCaps: [Capsule Dimension = 16, routings = 7]
Optimizer: ADAM (learningRate = 0.001)
Loss: binary_crossentropy (loss weight = [1. , 0.392])
Regularization Strength (Continual Learning): 1.7
Epochs: 100 (training) and 50 (training with Continual Learning)
Batch_Size = 20 (training) and 20 (training with Continual Learning)

Figure 3: CapsNet Architecture

by the weighted sum of all prediction vectors from the *PrimaryCapsules*. More specifically, the total input s_j (equation 2) to each digit capsule is the sum of these prediction vectors, each scaled by a coupling coefficient c_{ij}, which are learned through the dynamic routing process. The iterative dynamic routing process adjusts these c_{ij} so that if a prediction vector from a *PrimaryCapsules* agrees well with the actual output of a *DigiCaps* capsule, its corresponding c_{ij} is increased, meaning that it will contribute more to the total input of that higher-level capsule in the future [19].

$$s_j = \sum_i c_{ij}\hat{u}_{j|i}, \quad \hat{u}_{j|i} = W_{ij}u_i \quad (2)$$

To ensure the length of the capsule's output vector reflects the likelihood of the entity, a non-linear squashing function is employed [19]:

$$v_j = \frac{||s_j||^2}{1 + ||s_j||^2} \cdot \frac{s_j}{||s_j||} \quad (3)$$

where v_j represents the vector output of capsule j, and s_j represents its total input. The CapsNet's squashing function plays a crucial role in transforming the input vector s_j into an output vector v_j. This is to ensure the length of the output vector of a capsule is between 0 and 1 (but never reaching 1), where the length is meant to represent a probability, and probabilities range between 0 and 1. The function takes the total input to a capsule s_j and shrinks it to a small vector if it's short and to a vector of length just below 1 if it's long.

This transformation comprises several essential steps. To begin, the numerator $||s_j||^2$ calculates the square of the input vector's length, capturing the squared norm of s_j. In the denominator, $(1 + ||s_j||^2)$, the squared norm is adjusted by adding 1, ensuring that the output value remains within the range of 0 to 1. This adjustment serves as a vital normalization factor, effectively compressing the input vector's length. In the final step, $(s_j/||s_j||)$ scales the input vector by its original length, preserving its direction and orientation. This preservation of orientation is a crucial aspect of the squashing function, indicating the presence and characteristics of specific features or entities within the input. It plays a fundamental role in CapsNet's hierarchical feature representation. Since there are two classes, COVID and Non-COVID, the shape of DigiCaps is 2×16 (two 16-dimensional vectors). Each vector v_j functions as the capsule for class j. The probability of classification of an image is determined by $||v_j||$. Subsequently, two hidden fully connected layers with 628 and 1448 neurons are utilized to reconstruct the 300×300×1 image from $||v_j||$.

3.2.3 Vision Transformer (ViT)

Alexey Dosovitskiy *et al.* [3] first presented the Vision Transformer (ViT) model. For the purpose of classifying COVID-19 CT scan images, we follow the implementation of ViT as in [20]. The ViT model draws inspiration from the self-attention mechanism mostly applied in natural language processing (NLP) [6]. The ViT model uses a transformer architecture that uses attention to handle patch embeddings for image classification tasks.

Input images for ViT are decomposed into fixed number of non-overlapping square of patches. To apply ViT, we resize CT scan image of size 300×300 to 72×72 pixels and partition into 6×6 pixels to have 144 patches. Each partch is arranged into 1-dimensional vector similar to the flattening operation in CNN and linearly project into higher-dimentional embedding space to increase the feature learning capabilities of the model. These embeddings are augmented with positional encodings to maintain the patches' spatial relationships, forming a sequence similar to text in language models. This sequence is processed by the transformer's self-attention mechanism, allowing the integration of both local and global information across patches [3].

In the ViT model, there is a special part known as the 'class token' that is included along with the encoded sequences (image data). These tokens can be considered as placeholder or a tag that holds the final result after the model has analyzed the image. As the image data passes through the transformer's layers, which are a series of processing steps, this class token gathers important information that helps to determine what the image represents. After the data has moved through all the layers of the transformer, the class token has collected enough information to help decide the category of the image, such as whether it is an image of a Covid-19 positive or negative. ViT's innovative use of self-attention make the model to capture complex patterns for accurate image classification [2]. The configured hyperparameters for the ViT model comprises a learning rate of 0.001, 8 transformer layers, 4 multi-head attention heads, and a regularization strength of 2.0 to facilitate continual learning.

3.3 Exploring Transfer Learning to Generalize Models

Transfer learning involves reusing a model developed for one task as the starting point for another, particularly in neural networks where a pre-trained model on a general dataset is fine-tuned for specific tasks [25]. This approach leverages learned features to reduce training time and data needs. In the context of continual learning, transfer learning starts with a model pre-trained on a task T_0 with optimized parameters $\theta^*_{T_0}$. When a new task T_1 arises, the model adapts from $\theta^*_{T_0}$ to $\theta^*_{T_1}$, using knowledge from T_0 for better performance on T_1 efficiency. However, unlike pure transfer learning where the focus is solely on the current task, continual learning requires maintaining performance across all previous tasks. This necessitates strategies like regularization, weight freezing, or selective retraining to balance retaining knowledge from previous tasks while adapting to new ones, particularly to mitigate catastrophic forgetting

[17, 12]. As the model encounters more tasks $T_2, T_3, \ldots, T_n$, it incrementally learns and adapts, aiming to preserve performance on all prior tasks. Transfer learning, thus, is a vital strategy within the broader scope of continual learning, emphasizing efficient adaptation and knowledge retention, especially valuable in dynamic environments requiring models to continually adjust to new data and tasks.

We apply transfer learning technique to extend our model performance on the new dataset, particularly for the context of COVID-19 CT scan image classification. The ability of models to extend their learned knowledge to unseen tasks and datasets beyond their original training sets is extremely important, as outlined in [5] and [26]. Transfer learning is a technique using pre-trained weights from existing models to leverage shared features between prior and new tasks. Hence, we initially train our models using the *dataset_mod_dev* dataset and then apply the pre-trained weights to train models on the *dataset_mod_gen* dataset. As shown in the Figure 4 and Table 1, our models show effective generalization on the *dataset_mod_gen* through transfer learning. However, the models' performance significantly decrease on the original dataset (*dataset_mod_dev*) they were trained on. This issue, known as catastrophic learning, which is a crucial obstacle to develop robust deep learning models that produce better prediction accuracy on both original training COVID-19 CT dataset and new unseen COVID-19 CT dataset. Thus, we employ modified EWC regularization technique to address this issue.

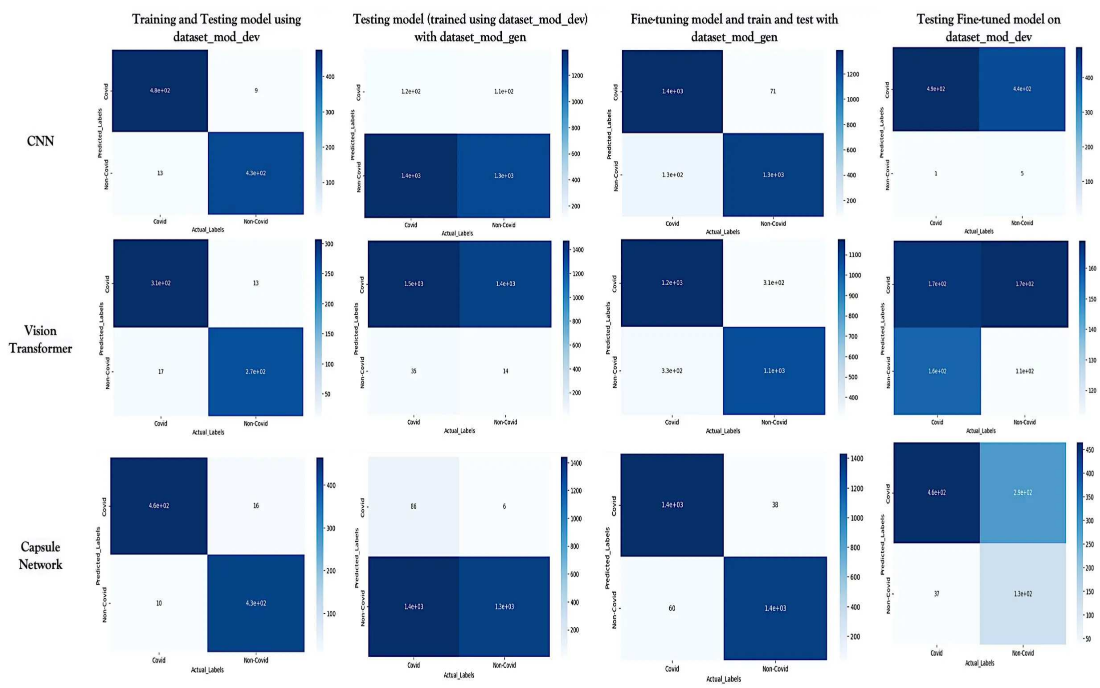


Figure 4: Confusion Matrix for Generalization Results Using Transfer Learning Technique

	CNN(COVID-CNN)				Vision Transformer				Capsule Network			
Tests Strategies	**Accuracy%**	**Precision%**	**F1 Score%**	**Recall%**	**Accuracy%**	**Precision%**	**F1 Score%**	**Recall%**	**Accuracy%**	**Precision%**	**F1 Score%**	**Recall%**
Training and Testing model using dataset_mod_dev	97.85	97.86	97.85	97.85	96.41	96.41	96.41	96.41	97.17	97.18	97.17	97.17
Testing model (trained using dataset_mod_dev) with dataset_mod_gen	47.96	48.3	36.87	47.96	51.38	40.66	0.362	51.38	49.75	72.26	36.2	49.76
Fine-tuning model and train and test with dataset_mod_gen	93.75	93.81	93.76	93.75	77.64	77.66	77.65	77.64	96.59	96.61	96.6	96.6
Testing Fine-tuned model on dataset_mod_dev	50.43	68.88	34.69	50.43	46.11	45.96	46.02	46.12	64.45	68.85	60.32	64.46

Table 1: Generalization Results Using Transfer Learning Technique

3.3.1 Learning Without Forgetting

In the field of continual learning, a significant challenge is mitigating catastrophic forgetting, which occurs when a neural network trained on a new task forgets the knowledge it has acquired from previous tasks. As shown in Table 1, even we use the transfer learning technique to achieve better prediction accuracy on the new *dataset_mod_gen* dataset, the performance of models decline on its original training *dataset_mod_dev* dataset. An effective approach to this challenge involves adjusting the learning objective for each new task so that retain information learned from previous tasks. We apply the similar idea proposed in [12] by adding a regularization term to the loss function that penalizes changes to those parameters that are crucial for previous tasks.

We select the categorical cross-entropy loss function for the objective function to train each model proposed in this paper: CNN, ViT, and CapsNet. The loss function is as follows:

$$\theta^*_{n+1} = \arg\min_{\theta} \left[\mathrm{CE}(y, \hat{y}(\theta)) + \lambda_i \sum_j \left(\theta_j - \theta^*_{i,j}\right)^2 \right] \quad (6)$$

Categorical Cross-entropy Loss (CE):

- $\mathrm{CE}(y, \hat{y}(\theta))$ represents the categorical cross-entropy loss.
  - y represents the true labels in the dataset.
  - $\hat{y}(\theta)$ represents the predicted probabilities by the three proposed models, parameterized by θ.

  This loss function measures the discrepancy between the true labels and the model's predictions, guiding the model to adjust its parameters to improve accuracy.
- The term $\sum_j \left(\theta_j - \theta^*_{i,j}\right)^2$ is the regularization term (a simplified version of the Elastic Weight Consolidation (EWC) in [12]) for each previous task T_i, where θ_j are the current weights of the model, and $\theta^*_{i,j}$ are the weights optimized for the task T_i. Thus, it measures the deviation of θ_j from $\theta^*_{i,j}$, the optimal parameters for task T_i which is the sum of squared differences between these sets of weights.
- λ_i is a weighting factor that balances the importance of the new task against the risk of

	CNN(*COVID-CNN*)				**Vision Transformer**				**Capsule Network**			
Tests Strategy	**Accu racy%**	**Prec ision%**	**F1 Score%**	**Re call%**	**Accu racy%**	**Prec ision%**	**F1 Score%**	**Re call%**	**Accu racy%**	**Prec ision%**	**F1 Score%**	**Re call%**
Training with Continual Learning using dataset_mod_gen and Testing the *Generalized Model* on dataset_mod_gen	79.33	79.83	79.28	79.33	77.05	77.09	77.06	77.05	80.35	80.40	80.29	80.35
Testing the *Generalized Model* (Transfer and Continual Learning) on dataset_mod_dev	82.26	83.12	82.00	82.26	81.96	81.95	81.96	81.86	79.35	83.52	79.35	79.08

Table 2: Generalization Results Using Transfer Learning With EWC

forgetting the previous tasks. The various values for the hyperparameter λ_i is tested based on the trial and error method. A value of 2.0 is found to optimally balance the acquisition of new knowledge from *dataset_mod_gen* dataset with the retention of previously learned information from *dataset_mod_dev* dataset.

We implement the proposed CNN model and ViT and CapsNet in a continual learning framework. The loss function' objective is to facilitate the acquisition of new knowledge related to the current task, as depicted by the categorical cross-entropy term, while also preserving the knowledge gained from previous tasks. It fine-tunes the model's parameters θ to decrease prediction errors for the current task. This loss function acts as a navigator, directing the model to achieve more accurate predictions and increased performance on new tasks. At the same time, the regularization term is crucial in protecting the knowledge of models that has already acquired. This component counteracts significant changes to the model's parameters $\theta^*_{i,j}$, which were previously optimized for earlier tasks. The severity of the penalty for deviating from these past parameters is governed by the λ_i values, allowing for the control of how closely the model adheres to its prior learning. This balancing act, achieved by minimizing the combined objective function, enables the models to excel in current tasks while retaining its expertise in past tasks. This approach addresses a key challenge in continual learning: finding the equilibrium between assimilating new information and retaining existing knowledge.

4 Experimental Results

4.1 Proposed Models Results

Leveraging transfer learning with continual learning, we apply pre-trained weights from the model trained on *dataset_mod_dev* to develop generalized models for *dataset_mod_gen* with EWC regularization technique. This method aims to balance the acquisition of new knowledge with the retention of previously learned information showing the unique performance attributes exhibited by models COVID-CNN, ViT, and CapsNet. The model performance for COVID-CNN, ViT and CapsNet are summarized in the Table 2 and Figure 5.

To evaluate model performance of the generalized models with EWC, we show the generalized model prediction accuracy on *dataset_mod_gen* dataset, and knowledge retention performance of the generalized models with EWC on *dataset_mod_dev* dataset. Our COVID-CNN demonstrates excellent performance across all metrics with 82.26% accuracy on *dataset_mod_dev* dataset and 79.33% prediction accuracy on *dataset_mod_gen*. ViT shows slightly lower 81.96% accuracy

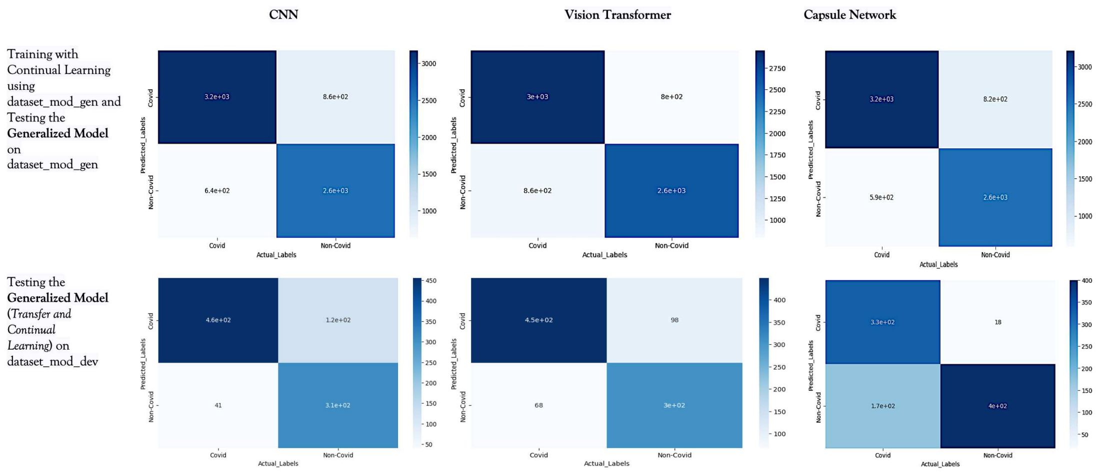


Figure 5: Confusion Matrix for Generalization Results Using Transfer Learning with Continual Learning

on *dataset_mod_dev* dataset and 77.05% prediction accuracy on *dataset_mod_gen* compared with COVID-CNN model. In contrast, CapsNet achieves 79.35% accuracy on *dataset_mod_dev* dataset and 80.35% prediction accuracy on *dataset_mod_gen* dataset. As a result, CapsNet performs best using transfer learning with continual learning in terms of accuracy, precision, F1 score, and recall. The COVID-CNN remains a reliable second choice due to its outstanding consistency and balance. Comparing the proposed models with EWC, COVID-CNN shows the best performance with balanced improvements across all metrics, ensuring reliable and accurate predictions on *dataset_mod_dev* dataset. ViT model shows lower prediction accuracy compared with COVID-CNN in overall consistency. The CapsNet shows least knowledge retention compared with the other models. Ultimately, COVID-CNN's consistent and balanced knowledge retention makes the model most effective on both datasets.

4.2 Comparison Results

To the best of our knowledge, our work is the first one using transfer learning with EWC technique for COVID-19 CT scan images analysis. We compare our results with the sate-of-the-art transfer learning models for the COVID-19 detection on CT scan images, such as Deep COVID DeteCT [13] and Deep-COVID [8]. These models utilized famous InceptionV3 [13] and NASNetLarge [8] pre-trained weights techniques and demonstrated substantial testing accuracy when trained on the *dataset_mod_dev* dataset. However, they show a significant decline in performance when evaluate on a novel *dataset_mod_gen* dataset. When comparing the Table 1 and Table 3, we find that our proposed COVID-CNN model, ViT and CapsNet outperforms the both Deep COVID DeteCT [13] and Deep-COVID [8] on *dataset_mod_dev* and *dataset_mod_gen* datasets. Furthermore, our models with EWC show more generalization ability compared with both Deep COVID DeteCT [13] and Deep-COVID [8] on *dataset_mod_gen* dataset.

	Deep-COVID[8]				Deep COVID DeteCT (DCD)[13]			
Test Strategy	**Accuracy%**	**Precision%**	**F1 Score%**	**Recall%**	**Accuracy%**	**Precision%**	**F1 Score%**	**Recall%**
Model training and testing using dataset_mod_dev dataset	95.59	95.61	95.59	95.59	87.63	87.77	87.65	87.63
Testing model on dataset_mod_gen dataset	28.19	28.29	28.15	28.19	41.1	41.06	41.07	41.1

Table 3: Generalization Results Using Existing Transfer Learning Models

5 Conclusion

In this paper, we develop three robust deep learning models with transfer learning and continual learning for the detection of COVID-19 in CT scan images. We find that the model generalization through transfer learning incurs catastrophic forgetting in our proposed CNN, and other ViT and CapsNet models. To address this, we implement a modified version of elastic weight consolidation (EWC) within our continual learning strategy, thereby alleviating the catastrophic forgetting associated with transfer learning. The original EWC technique is computationally intensive. With our simplified version of EWC, it takes less time and produces more accuracy during model training. Utilizing this approach, our developed COVID-CNN model, and the ViT and CapsNet show a great generalization on new *dataset_mod_gen* dataset, which were pretrained on *dataset_mod_dev* dataset. For our proposed models with EWC, COVID-CNN model shows superior performance over the ViT and comparable results to the CapsNet in adapting to new *dataset_mod_gen* dataset. Moreover, our COVID-CNN model achieves the highest knowledge retention rate of 82.26% for *dataset_mod_dev* dataset. The proficiency of COVID-CNN can be attributed to its compatibility with effective regularization, parameter efficiency, and a less complex architecture, fostering a balance between acquiring new knowledge and retaining existing information.

6 Acknowledgment

This work is supported by the National Science Foundation of the United States under Award 2200138. We thank anonymous reviewers for their insightful comments and inputs.

References

[1] A. A. Akinyelu and B. Bah. Covid-19 diagnosis in computerized tomography (ct) and x-ray scans using capsule neural network. *Diagnostics*, 13(8):1484, 2023.

[2] S. Cristina. The vision transformer model. https://machinelearningmastery.com/the-vision-transformer-model/, 2023. Accessed on 08 September 2023.

[3] A. Dosovitskiy, L. Beyer, A. Kolesnikov, D. Weissenborn, X. Zhai, T. Unterthiner, M. Dehghani, M. Minderer, G. Heigold, S. Gelly, and J. Uszkoreit. An image is worth 16x16 words: Transformers for image recognition at scale. *ArXiv Preprint ArXiv:2010.11929*, 2020.

[4] W. El-Shafai, A. Mahmoud, A. Ali, E. El-Rabaie, T. Taha, O. Zahran, A. El-Fishawy, N. Soliman, A. Alhussan, and F. Abd El-Samie. Deep cnn model for multimodal medical image denoising. *Comput. Mater. Contin*, 73:3795–3814, 2022.

[5] N. Enshaei, M. Rafiee, and F. Naderkhani. A generalization enhancement approach for deep learning segmentation models: Application in covid-19 lesion segmentation from chest ct slices. In *2022 30th European Signal Processing Conference (EUSIPCO)*, pages 1362–1366, 2022. https://api.semanticscholar.org/CorpusID:252997822.

[6] Z. Fu. Transformer in computer vision: Vit and its progress. https://api.semanticscholar.org/CorpusID:248986703, 2022.

[7] X. Gao, Y. Qian, and A. Gao. Covid-vit: Classification of covid-19 from 3d ct chest images based on vision transformer model. In *2022 3rd International Conference On Next Generation Computing Applications (NextComp)*, pages 1–4, 2021. https://api.semanticscholar.org/CorpusID:235731661.

[8] M. Ghaderzadeh, F. Asadi, R. Jafari, D. Bashash, H. Abolghasemi, and M. Aria. Deep convolutional neural network–based computer-aided detection system for covid-19 using multiple lung scans: Design and implementation study. *J Med Internet Res*, 23:e27468, 2021,4. http://www.ncbi.nlm.nih.gov/pubmed/33848973.

[9] S. Ghashghaei, D. Wood, E. Sadatshojaei, and M. Jalilpoor. Grayscale image statistical attributes effectively distinguish the severity of lung abnormalities in ct scan slices of covid-19 patients. *Springer Nature Applied Sciences*, 2023.

[10] S. Kato, M. Oda, K. Mori, A. Shimizu, Y. Otake, M. Hashimoto, T. Akashi, and K. Hotta. Classification and visual explanation for covid-19 pneumonia from ct images using triple learning. *Scientific Reports*, 12:20840, 2022.

[11] S. Khan, I. Ullah, N. Ullah, S. Shah, M. Affendi, and B. Lee. A novel ct image de-noising and fusion based deep learning network to screen for disease (covid-19). *Scientific Reports*, 13:6601, 2023.

[12] J. Kirkpatrick, R. Pascanu, N. Rabinowitz, J. Veness, G. Desjardins, A. Rusu, K. Milan, J. Quan, T. Ramalho, A. Grabska-Barwinska, and D. Hassabis. Overcoming catastrophic forgetting in neural networks. *Proceedings Of The National Academy Of Sciences*, 114:3521–3526, 2017.

[13] E. Lee, J. Zheng, E. Colak, M. Mohammadzadeh, G. Houshmand, N. Bevins, F. Kitamura, E. Altinmakas, E. Reis, J. Kim, and C. Klochko. Deep covid detect: an international experience on covid-19 lung detection and prognosis using chest ct. *NPJ Digital Medicine*, 4:11, 2021.

[14] Y. Liu, H. Li, T. Luo, C. Zhang, Z. Xiao, Y. Wei, Y. Gao, F. Shi, F. Shan, and D. Shen. Structural attention graph neural network for diagnosis and prediction of covid-19 severity. *IEEE Transactions On Medical Imaging*, 42:557–567, 2022.

[15] Maede Maftouni, Andrew Chung Chee Law, Bo Shen, Zhenyu James Kong Grado, Yangze Zhou, and Niloofar Ayoobi Yazdi. A robust ensemble-deep learning model for covid-19 diagnosis based on an integrated ct scan images database. In *IIE annual conference. Proceedings*, pages 632–637. Institute of Industrial and Systems Engineers (IISE), 2021.

[16] H. Miao and F. Tong. Convolutional neural network models of neuronal responses in macaque v1 reveal limited non-linear processing. *bioRxiv*, 2023.

[17] Francesco Pelosin. Dissecting continual learning a structural and data analysis, 2023.

[18] Mats L Richter, Julius Schöning, Anna Wiedenroth, and Ulf Krumnack. Should you go deeper? optimizing convolutional neural network architectures without training. In *2021 20th IEEE International Conference on Machine Learning and Applications (ICMLA)*, pages 964–971. IEEE, 2021.

[19] S. Sabour, N. Frosst, and G. Hinton. Dynamic routing between capsules. In *Advances In Neural Information Processing Systems*, volume 30, 2017. https://proceedings.neurips.cc/paper_files/paper/2017/file/2cad8fa47bbef282badbb8de5374b894-Paper.pdf.

[20] K. Salama. Keras documentation: Image classification with vision transformer. https://keras.io/examples/vision/image_classification_with_vision_transformer/, 2021. Accessed on 15 July 2023.

[21] Humam Adnan Sameer, A. H. Mutlag, and S. Gharghan. Deep learning-based covid-19 diagnosis using ct scans with laboratory and physiological parameters. *IET Image Processing*, Preprint, 2023.

[22] E. Soares, P. Angelov, S. Biaso, M. Froes, and D. Abe. Sars-cov-2 ct-scan dataset: A large dataset of real patients ct scans for sars-cov-2 identification, 2020. https://www.medrxiv.org/content/

early/2020/05/14/2020.04.24.20078584.

[23] K. Uddin, S. Dey, H. Babu, R. Mostafiz, S. Uddin, W. Shoombuatong, and M. Moni. Feature fusion based vggfusionnet model to detect covid-19 patients utilizing computed tomography scan images. *Scientific Reports*, 12:21796, 2022.

[24] B. Xiao, Z. Yang, X. Qiu, J. Xiao, G. Wang, W. Zeng, W. Li, Y. Nian, and W. Chen. Pam-densenet: A deep convolutional neural network for computer-aided covid-19 diagnosis. *IEEE Transactions On Cybernetics*, 52:12163–12174, 2021.

[25] D. Yang, C. Martinez, L. Visuña, H. Khandhar, C. Bhatt, and J. Carretero. Detection and analysis of covid-19 in medical images using deep learning techniques. *Scientific Reports*, 11:19638, 2021.

[26] J. Yang, A. Soltan, and D. Clifton. Machine learning generalizability across healthcare settings: insights from multi-site covid-19 screening. *NPJ Digital Medicine*, 5, 2022. https://api.semanticscholar.org/CorpusID:246702315.

[27] L. Zhou, X. Meng, Y. Huang, K. Kang, J. Zhou, Y. Chu, H. Li, D. Xie, J. Zhang, W. Yang, and N. Bai. An interpretable deep learning workflow for discovering subvisual abnormalities in ct scans of covid-19 inpatients and survivors. *Nature Machine Intelligence*, 4:494–503, 2022.