Int. J. Simul. Multidisci. Des. Optim. **17**, 15 (2026)

https://doi.org/10.1051/smdo/2026014

IJSMDO
Available online at:
https://www.ijsmdo.org

AI-Driven Simulation and Neural Optimization for Smart Systems and Healthcare Applic

RESEARCH ARTICLE OPEN ACCESS

Adaptive hyperparameter and regularization strategies for improved elastic weight consolidation in lung cancer classification

Zainab Muhammed[1,2] and Belal Al-Khateeb[3,*]

[1] Informatics Institute for Postgraduate Studies, Information Technology & Communication University, 10074, Baghdad, Iraq
[2] Computer Science Department, College of Education for Pure Science/ Ibn-Al Haitham, University of Baghdad, 10069, Baghdad, Iraq
[3] College of Computer Science and Information Technology, University of Anbar, 10003, Ramadi, Iraq

Received: 2 July 2025 / Accepted: 24 June 2026

Abstract. Accurate and reliable categorization of lung cancer histopathology is considered challenging due to incremental data arrival, distributional shifts, and the tendency of conventional convolutional neural networks to suffer from overfitting and catastrophic forgetting when new subclasses are introduced. To address these issues, a continual learning framework was developed to preserve previously acquired knowledge while maintaining high diagnostic performance. A ResNet50 backbone regularized with Elastic Weight Consolidation (EWC) was employed, in which the Fisher-weighted penalty coefficient (λ) was updated dynamically through a curvature-aware schedule. By doing so, stability and plasticity were balanced adaptively without the need for manual grid search. The framework was trained and evaluated on 15,000 pathologically confirmed lung tissue images comprising adenocarcinoma (aca), squamous cell carcinoma (scc), and benign (n) classes. The experimental setup was organized as a continual learning sequence over three phases of training, where new classes were introduced incrementally while earlier ones were revisited. The proposed approach achieved external test accuracy of 100%, 97.2%, and 98.96% across the three phases, demonstrating its capability to sustain high accuracy while effectively preventing catastrophic forgetting. These results indicate that the integration of adaptive hyperparameter tuning with a curvature-driven EWC schedule can enhance continual learning performance in lung cancer classification. It is concluded that the ResNet50-EWC framework provides a scalable and clinically relevant solution that eliminates the need for full retraining when data expands, and future validation on additional histological subtypes and external whole-slide cohorts is planned.

Keywords: Elastic weight consolidation (EWC) / adaptive hyperparameters / regularization strategies / catastrophic forgetting / continual learning / neural network optimization

1 Introduction

Lung cancer remains the leading cause of cancer-related mortality worldwide, accounting for an estimated 1.8 million deaths in 2023 alone [1]. Early and accurate histopathological diagnosis is pivotal, because therapeutic decisions and prognostic assessments depend on correctly identifying tumor subtypes, most commonly adenocarcinoma, squamous-cell carcinoma, and large cell carcinoma according to the latest WHO classification [2]. Traditional light-microscopy workflows, however, are labor-intensive, prone to inter-observer variability, and increasingly strained by the growing volume of biopsies generated through low-dose CT screening programs [3]. These challenges have catalyzed the adoption of computational pathology and deep-learning approaches that can support or partially automate slide interpretation [4,5].

Convolutional neural networks (CNNs) have demonstrated near-human performance in diverse image-analysis tasks, including mitotic-figure detection [6], breast-cancer grading [7], and colorectal polyp triage [8]. In lung-histopathology, high-capacity backbones such as VGG-19, Inception, and ResNet-50 routinely exceed 95% classification accuracy on benchmark datasets like LC25000 [9]. Despite these advances, three limitations persist. First, most published pipelines are trained once and then froze; when new staining protocols or rare tumor subtypes are introduced, the entire network must be retrained from scratch, incurring prohibitive computational costs and risking catastrophic forgetting of previously learned classes [10]. Second, hyper-parameters (e.g. learning rate,

* e-mail: belal-alkhateeb@uoanbar.edu.iq

regularization strength) are almost invariably selected by grid search, an offline procedure that scales poorly and is sensitive to dataset-specific noise. Third, many studies report single-split results without external validation, leading to optimistic performance estimates and limited clinical utility.

Continual-learning algorithms attempt to mitigate catastrophic forgetting by encouraging parameter stability across tasks [11]. Elastic Weight Consolidation (EWC) is one of the most widely adopted strategies; it adds a Fisher-information-weighted quadratic penalty that keeps parameters close to values important for past tasks. While effective, classical EWC relies on a fixed regularization coefficient, whose best value varies with data complexity and class imbalance. Recent work in natural-image recognition suggests that dynamically adjusting λ based on loss-curvature estimates can improve stability–plasticity trade-offs while reducing hyper-parameter tuning overhead [12].

Although deep-learning approaches have demonstrated strong performance in histopathology, their integration into routine clinical workflows remains constrained by several limitations. Existing models often suffer from catastrophic forgetting when retrained with novel tissue subtypes or staining protocols. In addition, hyperparameter optimization is predominantly performed through exhaustive grid search, a process that is computationally expensive, inefficient, and difficult to scale. The limited adaptability of current methods to evolving datasets, coupled with insufficient external validation, further undermines their robustness and generalizability across heterogeneous clinical settings.

The main problems that are identified in current methodologies are catastrophic forgetting, inefficient hyperparameter optimization, and limited adaptability to evolving datasets. To address these challenges, this study proposes a ResNet50-based framework enhanced with EWC and an adaptive hyperparameter scheduler. The incorporation of EWC enables continual learning without catastrophic degradation of prior knowledge, while the adaptive scheduler dynamically adjusts hyperparameters, eliminating the need for costly manual tuning. Together, these innovations reduce computational overhead, improve stability, and enhance the clinical applicability of deep-learning models in dynamic and evolving environments.

In this paper, three main contributions are presented:

- A ResNet50 backbone is augmented with a curvature-aware variant of EWC, enabling continual learning without catastrophic forgetting of previously acquired knowledge.
- A dynamic hyperparameter scheduler is proposed, which automatically adjusts λ based on the second-order statistics of the validation loss, thereby eliminating the need for manual grid search.
- The structure of the EWC framework is adapted to support realistic continual learning scenarios in which the number of classes is not fixed and may vary dynamically across training phases.

The paper is organized as follows in the remaining sections. In Section 2, the related works are presented, while Section 3 introduces the proposed method. Section 4 presents the results and discussions, and Section 5 shows the limitations of the study, while Section 6 concludes with a few suggestions for future research.

2 Related works

In the rapidly evolving domain of medical image analysis, deep learning approaches, particularly convolutional neural networks, have demonstrated considerable promise in the identification of lung cancer and the improvement of diagnostic imaging. Numerous foundational studies have leveraged extensive image datasets, with considerable research focusing on large-scale collections of histopathology images to train and validate sophisticated CNN architecture [13]. This section will discuss prior research in the domain of deep learning and the application of continuous deep learning techniques in the classification of lung cancer.

The study in [14] involved the creation of a specialized nine-layer CNN designed to categorize lung tissue into (benign, adenocarcinoma, and squamous cell carcinoma) utilizing the LC25000 dataset. The model, trained on histopathology patches measuring 180×180 pixels, attained a validation accuracy of 97.2%, with a training accuracy of 96.11%. The model relied solely on a single internal train/test split, excluding the implementation of cross-validation or validation against an external dataset. The limitations of this approach present considerable issues related to overfitting and the capacity to generalize results. Moreover, the architecture of CNN demonstrates a certain level of simplicity, and the paper lacks analyses related to interpretability or approaches for optimization in real-world applications.

The research presented in [15] investigates the application of a shallow CNN architecture for the classification of histopathological images derived from the LC25000 dataset. The model specifically targets three classes of lung cancer: benign, adenocarcinoma, and squamous cell carcinoma, along with two classes of colon cancer. The CNN is designed with three layers of convolution and pooling, specifically optimized to train on image patches while reducing sampling to preserve texture details. The model exhibited an accuracy rate exceeding 97% in the classification of lung cancer and surpassed 96% in the classification of colon cancer. The limitations of the research include a lack of detailed information on cross-validation, and the failure to incorporate external datasets for evaluation raise concerns about potential overfitting and limited generalizability. Moreover, the simplicity of the model architecture and the ambiguity surrounding dataset partitioning methods may impact its dependability in clinical environments.

The research explained in [16] an ensemble CNN model aimed at classifying three categories of lung tissue: normal, adenocarcinoma, and squamous cell carcinoma, utilizing the LC25000 dataset. The model incorporates Grad-CAM to enhance explainability by visually emphasizing critical diagnostic areas within the images. This methodology exhibited remarkable classification performance, achieving accuracy levels between 97.11% and 99.69%, with AUC

Algorithm 1. Training with EWC Regularization.

Input:
D_t: Training data for task t
θ: Model parameters (initialized from previous task)
F: Fisher information matrix from previous tasks
λ_0, $\Delta\lambda$: Initial λ value and step size
η_0, decay: Initial learning rate and decay factor
Output:
Updated parameters θ
Stored θ* and F for the next task
Steps:
1. For each epoch 1 ... MaxEpoch:
2. For each minibatch (x, y) ∈ D_t:
3. Compute forward pass: logits ← Model(x; θ)
4. Compute task loss: L_{task} ← FocalLoss(y, logits)
5. If $t > 1$, compute EWC penalty: $L_{EWC} \leftarrow \Sigma_i \frac{1}{2} F_i (\theta_i - \theta^*_i)$
6. Else set $L_{EWC} = 0$
7. Compute total loss: $L_{total} = L_{task} + \lambda \cdot L_{EWC}$
8. Update model parameters:
9. a. Compute gradients: $g = \nabla\theta\ L_{total}$
10. b. Update parameters: $\theta \leftarrow \theta - \eta \cdot g$
11. Update λ based on validation performance:
12. If validation accuracy decreases: $\lambda \leftarrow \lambda + \Delta\lambda$
13. Else: $\lambda \leftarrow \max(\lambda - \Delta\lambda, \lambda_0)$
14. Update learning rate: $\eta \leftarrow \eta_0 / (1 + \text{decay} \times \text{epoch})$
15. Return updated θ, and store θ* and F for the next task.

values ranging from 99.77% to 99.94%. The limitations of this study include the lack of external or cross-dataset validation, which raises concerns about overfitting or potential data leakage, as well as the increased computational costs associated with ensemble learning and visualization techniques.

The research presented in [17] improved VGG-19 architecture that incorporates residual connections in conjunction with preprocessing methods, including adaptive histogram equalization and CLAHE. The model was evaluated using the three-class lung cancer subset of the LC25000 dataset, which includes adenocarcinoma, squamous cell carcinoma, and benign cases, resulting in a classification accuracy of 97.73%. The model demonstrates robust performance; its efficacy is significantly reliant on image augmentation techniques. The limitation of this study is the absence of validation on external datasets or whole-slide images introduces concerns regarding its generalizability in clinical applications. Furthermore, the incorporation of residual blocks elevates the computational complexity in comparison to the conventional VGG-19 architecture.

The research in [18] demonstrated the identification of lung cancer through advanced deep learning methodologies, particularly emphasizing the function of CNNs. It utilized these techniques for the automated classification and prediction of lung cancer in histopathological images, thereby minimizing dependence on manual evaluation. The study employs an extensive array of histopathological images pertinent to lung cancer. The CNN architecture is designed to extract features and recognize patterns associated with lung cancer. The model employed in the cited research attained a score of 95%, indicating its effectiveness in precise classification. The results demonstrate that the model significantly improves the early identification of lung cancer. The constraints identified in the study underscore the significant impact that advanced computational approaches have on improving the precision and effectiveness of diagnostic processes conducted by pathologists. The manuscript lacks adequate detail regarding data preprocessing, augmentation techniques, and computational requirements, all of which are critical factors influencing model performance and scalability. The identified factors underscore the necessity for additional validation and refinement of the methodology in subsequent research endeavors.

Another study in [19] explored the utilization of ResNet50 architecture. The ResNet50 model has been engineered to classify lung cancer into distinct categories using histopathological images. The model exhibited an impressive accuracy of 98%, indicating its potential for precise diagnosis. The study drawback specific constraints, it is crucial to conduct external validation across various datasets to establish their relevance in practical situations.

In our previous study [20], we present a continual learning framework that integrates a conventional convolutional neural network (CNN) with Elastic Weight Consolidation (EWC), referred to as CNN-EWC, to address the issue of catastrophic forgetting during sequential training. EWC incorporates a regularization term that imposes penalties on substantial modifications to parameters considered essential for previously acquired tasks, thus safeguarding existing knowledge. The model is subjected to continuous training across three balanced histopathological classification tasks, which include a dataset of 15,000 images that depict benign lung tissue, adenocarcinoma, and squamous cell carcinoma. CNN-EWC demonstrates sustained test accuracy between 98% and 99.6% across different task sizes, specifically ranging from 3,000 to 6,000 images, without the need for additional retraining. Nonetheless, a limitation of this study is that each training phase is constrained to a fixed number of classes, specifically three, without altering the number of classes across different tasks. The uniform task structure may not adequately represent the intricacies of real-world continual learning scenarios, in which the quantity and diversity of classes can evolve over time.

The research presented in [21] employed continuous learning methodologies to improve forecasting accuracy within regional energy markets. This was achieved through the utilization of a dataset that included 1 yr of 15 min resolution data sourced from ten renewable energy generators and ten local consumer facilities. The research encompassed methodologies including Learning Without Forgetting (LWF), Elastic Weight Consolidation (EWC), Online-EWC, and Synaptic Intelligence (SI). The conclusions indicated that both Online-EWC and EWC effectively mitigated catastrophic forgetting, leading to a notable decrease in its occurrence. The trials were conducted using data from a small power grid, which represents a limitation of the study. The findings presented

in the article are applicable to various levels of the power grid where catastrophic forgetting may take place. The limitation of paper is the phenomenon of catastrophic forgetting may occur. The research is confined to a particular regional energy market, which could influence the applicability of the findings to broader contexts. The research is confined to a particular regional energy market, which could influence the applicability of the findings to broader contexts. The analysis does not comprehensively consider sustained performance in the context of concept the complexities associated with real-time implementation. Furthermore, the aspects of interpretability and computational efficiency concerning continuous learning models are not thoroughly examined.

The research in [22] presented the investigation of the utilization of continual learning techniques to improve domain adaptation in the classification of chest X-rays. Their assessment involves methodologies including Joint Training (JT), EWC, and LWF, utilizing the ChestX-ray14 and MIMIC-CXR dataset for analysis. The results demonstrated that JT attains superior overall performance in both source and target domains. LWF exhibits competitive outcomes while not necessitating access to source domain data, underscoring its applicability in situations where data privacy considerations are paramount. The drawback of paper is the risk of catastrophic forgetting in specific continual learning methodology and the difficulties related to processing across varied medical image datasets.

In summary, previous studies have shown the effectiveness of deep learning and CNN-based architectures in classifying histopathological images of lung cancer. Nonetheless, most of these studies are constrained by the absence of external validation, dependence on single train/test splits, limited class settings, and inadequate attention to catastrophic forgetting in dynamic contexts. Furthermore, many methodologies have concentrated exclusively on static learning frameworks, neglecting the critical need for ongoing adaptation in response to changing datasets and clinical settings.

The identified research gaps highlight the necessity for models that extend beyond static classification and can maintain performance in continual learning environments. There is a notable deficiency in strategies that guarantee knowledge retention while accommodating new tasks, incorporate interpretability to foster clinical trust, and uphold computational efficiency appropriate for large-scale implementation. Furthermore, the lack of extensive validation on diverse and external datasets further limits the applicability of previous methods. The identified gaps provide the impetus for the proposed method, which aims to enhance both performance and memory protection in practical clinical scenarios.

3 Proposed model

The model depicted in Figure 1 is structured to proficiently acquire new tasks while reducing the likelihood of catastrophic forgetting a prevalent issue in continuous learning contexts. To resolve this issue, the model utilizes the Fisher Information Matrix to identify the parameters that are most critical to tasks that have been learned previously. To prevent any changes to those important parameters while training is ongoing, a regularization penalty is then applied. Moreover, the methodology incorporates carefully selected hyperparameters enabling accurate control over the model's learning processes. This adaptability promotes a better balance between maintaining existing knowledge and incorporating new information.

3.1 Input

Loading dataset from [23], which contains five classes of histopathology images (colon and lung cancer), only three classes for lung cancer are used. There are 5000 images of Lung benign tissue(n), 5000 images of Lung adenocarcinoma (aca), and 5000 images of Lung squamous cell carcinoma (scc), every image in the dataset is in jpeg format and has a resolution of (768×768) pixels.

The initial phase of the process entails the input of histopathology images, which are classified into two classes:

- Benign (n): Denoting healthy or non-malignant tissue specimens.
- Adenocarcinoma (aca): Denoting malignant or neoplastic tissue specimens.

This classification framework supports a supervised learning approach. The model was trained to distinguish between normal and abnormal (cancerous) tissue using labeled image data. The annotated examples provide the fundamental ground truth necessary to guide the learning process and improve model performance.

3.2 Preprocessing

To improve performance and speed up for approach convergence, the input images are subjected to a preparation pipeline that aims to standardize and improve data quality:

- Resizing: Each input image undergoes a transformation to achieve a standardized dimension of (224×224) pixels. This procedure ensures alignment with the input specifications of prevalent deep learning frameworks, thereby mitigating shape-related discrepancies throughout the training process.
- Normalization: The pixel intensity values undergo a scaling process to fit within a range of 0 to 1. The normalization process facilitates faster convergence and improves training stability by maintaining the input values within a comparable numerical range.

3.3 Setup hyperparameter

The careful selection and configuration of hyperparameters are crucial in the development of deep learning models, significantly influencing their performance and generalization capabilities. Hyperparameters, in contrast to model parameters that are acquired during the training phase, are predetermined settings established prior to the training

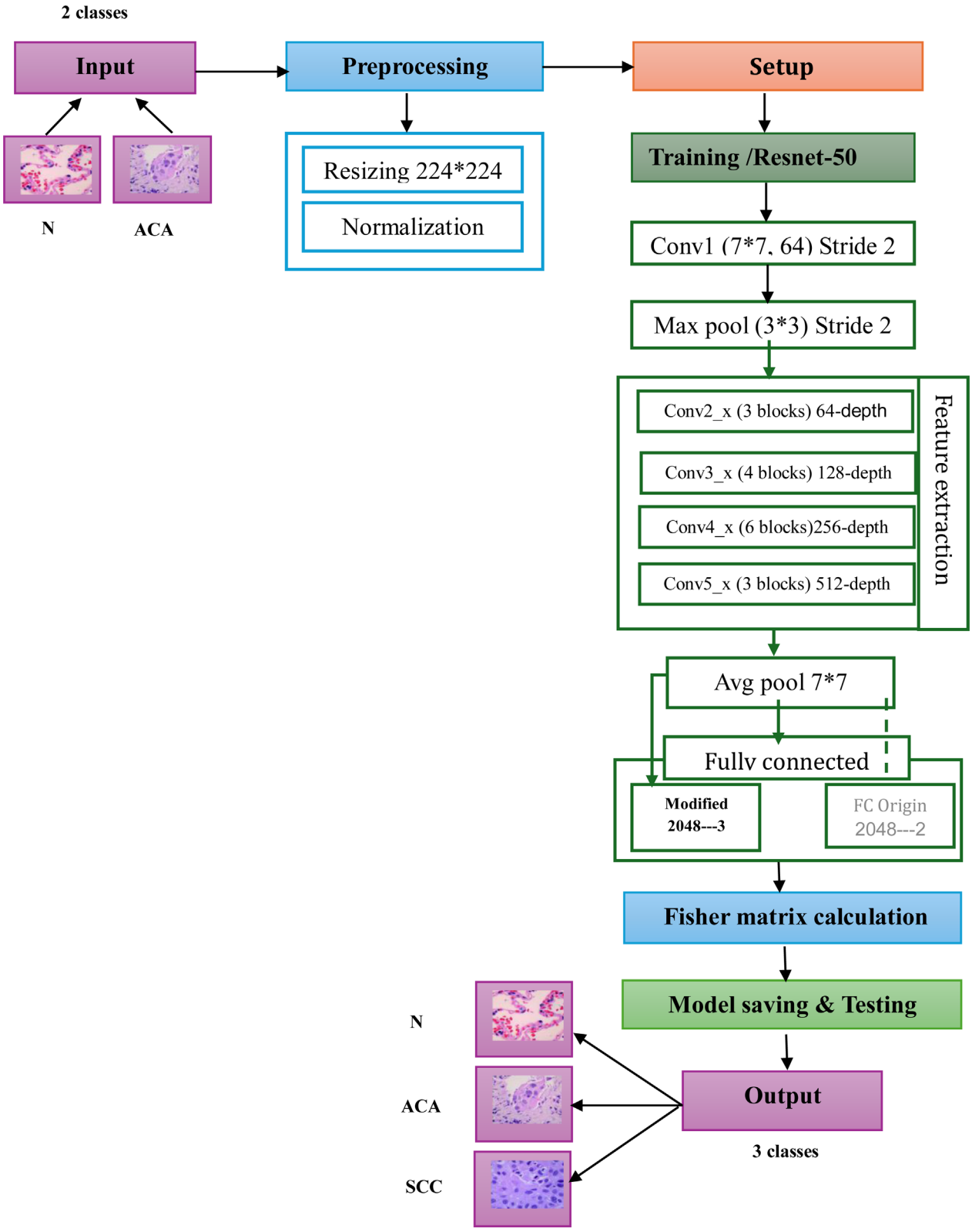

Fig. 1. The proposed approach (Resnet50-EWC).

Table 1. Hyperparameter of proposed approach.

Hyperparameter	Value(s)	Justification
Epochs	20 epochs with early stop	Ensures sufficient training; early stopping avoids overfitting and halts training once performance plateaus.
Learning Rate	Task 1: 0.00002 Task 2: 0.0001 to 0.000025 Task 3: 0.0002 to 0.0001	Stable convergence for Task 1; Task 2 and Task 3 uses a dynamic range to optimize learning and prevent destabilization.
EWC Lambda	Task 1: Not used Task 2: 14863.87 to 10308.57 Task 3: 14863.87 to 10308.5	Used only in Task 2 and Task 3 to prevent catastrophic forgetting; higher values indicate strong regularization of previous knowledge.
EWC Penalty (γ)	Task 1: Not used Task 2: 743.19 to 515.43 Task 3: 743.19 to 515.43	Applies regularization strength in EWC; helps balance plasticity and stability in Task 2.
Dropout Rate	0.25	Helps prevent overfitting by randomly deactivating neurons during training.
Batch Size	16	A small batch size allows for more updates per epoch and can improve generalization.
Training Size of Data	80%	A large portion of data ensures effective model learning.
Testing Size of Data	20%	Reserved to evaluate model performance and detect overfitting.
Loss Function	Cross-Entropy Loss	Commonly used for classification tasks; measures the performance of output probabilities.
Optimizer	ADAMW	Combines the benefits of Adam with weight decay for better generalization.
Activation Function	ReLU	Efficient and widely used; introduces non-linearity and avoids vanishing gradients.

process and influence various facets of the learning dynamics. This research presents a systematic methodology for hyperparameter optimization across two consecutive tasks, with the objective of enhancing model efficacy and preserving knowledge retention.

The selection of critical hyperparameters, including the number of epochs, learning rate, batch size, dropout rate, and optimizer choice, significantly impacts the model's convergence behavior, training stability, and generalization ability. Furthermore, in the context of continual learning, particular parameters such as EWC Lambda and Penalty (γ) play a vital role in alleviating catastrophic forgetting. These parameters implement regularization mechanisms that help maintain previously acquired knowledge.

This configuration utilizes a task-specific methodology wherein hyperparameters are dynamically adjusted in response to the demands of the training phase. Table 1 illustrates Task 1 highlights the convergence of the initial model, while Task 2 integrates mechanisms like early stopping and EWC to improve the learning process and protect against the loss of previously acquired knowledge. This comprehensive hyperparameter strategy ensures a balanced learning approach that satisfies the requirements of both stability and plasticity in sequential tasks.

3.4 Training

The approach illustrated above utilizes the (ResNet-50 model) that considered the backbone for feature extraction. The architecture initiates with a convolutional layer measuring 7×7, incorporating 64 filters and a stride 2. This is succeeded by a max pooling layer of dimensions 3×3, which similarly employs a stride 2. The preliminary layers serve to diminish spatial dimensions while extracting fundamental features. The fundamental structure of ResNet-50 is composed of four sequential convolutional blocks, labeled $Conv2_x$ through $Conv5_x$, which exhibit an increase in depth from 64 to 512. This architecture is designed to capture progressively more complex and abstract features from the input images. Following the feature extraction phase, a global average pooling operation utilizing a 7×7 kernel is implemented to reduce the spatial feature maps into a concise vector appropriate for classification purposes.

In Task 1, the original fully connected (FC) layer of ResNet-50 is retained, which maps the extracted 2048-dimensional feature vector to two output classes "n" and ″aca″. For Task 2, however, the classification task is expanded to include an additional class, Squamous Cell Carcinoma (SCC). As a result, the FC layer is modified to accommodate a new output dimensionality. The feature vector is first extended or reconfigured to a size of 2408, followed by an updated FC layer that outputs three classes "n", "aca", and "scc". This adjustment enables the model to adapt to a more complex classification problem while reusing the foundational feature extraction learned during Task 1.

The model is trained in new tasks without losing performance on old tasks, the incorporation of the EWC penalty into the training process. Algorithm 1 depicted the step of training with EWC regularization.

Training Steps in Each Epoch Using EWC:

– Forward Pass: Images are passed through the model to produce predictions (logits). Sometimes the logits are adjusted using temperature to smooth the values.
– Task Loss Calculation: The improved Focal loss is used to calculate the classification error. Loss is denoted by:

$$L_{task}$$

– EWC Penalty Calculation (if applicable): This is done only after the first task. Equation (1) can be expressed as:

$$L_{EWC} = \sum_i F_i(\theta_i - \theta_i^*)^2 \tag{1}$$

Where L_{EWC}: is the total EWC loss (a penalty added to the task loss to protect past knowledge), $\sum_i$: is summation over all model parameters (indexed by iii), F_i: is fisher Information for parameters estimating how important this parameter is to the previous task. Higher means more important, θ_i: is the current value of parameter iii (during training on the new task), θ_i^*: is the old value of parameter iii, saved after training on the previous task, and $(\theta_i - \theta_i^*)^2 \, is \, s$quared difference that measures how much the parameter has changed. Larger changes mean greater deviation from previous knowledge.

– Combining the two losses: The final loss is calculated using this formula in equation (2):

$$L_{total} = L_{task} + \lambda L_{EWC} \tag{2}$$

Where $L_{total} \, is \, t$otal loss used for backpropagation (It combines task performance with memory preservation), L_{task} is task loss (the standard loss (e.g., Cross-Entropy or Focal Loss) for the current task. It measures how well the model is performing on the new data), L_{EWC} is EWC penalty (the regularization term from EWC that penalizes changes to important weights from previous tasks), and λ Lambda is a hyperparameter that controls the strength of the EWC penalty (a higher λ\lambda means more focus on preserving old knowledge; a lower one focuses more on learning the new task).

– Backward scrolling and weight updating.

EWC regularization prevents the model from forgetting what it learned in previous tasks by preventing important weights from changing too much. It adds ″memory″ to the model and enables it to learn continuously without starting from scratch for each task.

3.5 Fisher matrix calculation

The training steps are used to update the EWC information after each learning task, so that the model can later learn new tasks without forgetting the old ones.

– Model processes task-specific data: The model processes the current task data, i.e., the task data (e.g., images) is passed through the model to obtain results and implementations, in preparation for calculating the importance of each weight.
– Squared gradients accumulate into Fisher matrix: Fisher matrix accumulation that calculates the gradient of each weight relative to the loss, then square these gradients and sum them to estimate the importance of each weight to the current task. This produces what is known as a Fisher matrix, which represents the "sensitivity" of each weight. Equation (3) can be expressed as:

$$F_i \approx E\left[\left(\frac{\partial L}{\partial \theta_i}\right)^2\right] \tag{3}$$

Where (F_i:) is Fisher Information for parameter (θ_i) (measures parameter importance., (≈) Indicates this is an approximation, (E [·]) Expectation, average over data samples, ($\partial L/\partial \theta_i$) is Gradient of the loss with respect to parameter θ_i, and $(\partial L/\partial \theta_i)^2$ is squared gradient. If F_i is large changing θ_i causes a big change in the loss and If F_i is small the parameter does not affect the loss much and can be changed more freely.

– Matrix normalized and combined with prior Fisher matrices: Matrix normalized and combined with prior Fisher matrices. After calculating the new Fisher matrix, it is normalized (e.g., to downscale or balance), then combined with the matrices specific to the previous tasks. The goal is to accumulate knowledge about the importance of weights across tasks. Equation (4) can be expressed as:

$$\text{Updated} \quad F_i = (\text{decay} \quad \text{factor}) \times F_i^{old} + F_i^{new} \tag{4}$$

Where (F_i) is revised fisher information for parameter θ_i (reflecting cumulative significance across tasks), (Decay factor) introduces a curvature-driven adaptation of the Fisher matrix. Standard EWC assumes that all previously computed Fishers should be preserved equally, which may overemphasize outdated curvature information as tasks accumulate. Our rule allows the Fisher matrix to gradually forget older curvature estimates, thereby reducing the risk of over-constraining parameters and improving plasticity when facing new tasks. This results in a more balanced trade-off between stability retaining old knowledge and plasticity (adapting to new tasks). It mitigates catastrophic interference in long task sequences where rigid accumulation would otherwise degrade performance. The decay factor α\alphaα was empirically selected through validation, ensuring a balanced trade-off between stability and plasticity, and sensitivity analysis confirmed that performance remains robust across a reasonable range of values, (F_i^{old}) is the fisher information derived from prior tasks (indicating historical significance), and (F_i^{new}) is the newly computed Fisher Information derived from the current task, signifies its present significance.

– Current parameters stored for future reference: Storing current weight values for later use. A copy of the current weights is saved.

3.6 Model saving and testing

Upon completing the Fisher computation, the model saves its states and evaluates performance. This step is crucial to detect catastrophic forgetting and to determine whether

further intervention (e.g., rehearsal) is required. Several tests were conducted to verify the validity of the model and its ability to retain important information from each task. This ensures that the model can effectively apply prior knowledge when a new task is introduced.

The model was first evaluated using a set of images from the initial task, after training on two distinct tasks conducted at separate intervals. In addition, we performed an extra evaluation with image sets of varying sizes that were not included in the training data, to further assess classification accuracy.

Finally, the performance of the proposed model was assessed using standard metrics, including accuracy, precision, recall, and F1-score.

3.7 Output

In the final stage illustrated in Figure 1, the modified ResNet-50 receives the pre-processed histopathological patches and produces a three-dimensional output vector corresponding to the classes' "aca", "n",and "scc". Although the original dataset was dichotomous, the network's classification capacity was extended by replacing the ImageNet-specific fully connected layer with a new dense layer consisting of three units, followed by a SoftMax activation to yield class-conditional probabilities. During inference, the class with the highest posterior probability is reported as the prediction. After convergence, the trained weights are serialized "Model saving" and subsequently employed to evaluate unseen slides "Testing", thereby quantifying the network's generalization performance across the newly introduced third category.

4 Results and discussion

The results of the proposed model, which utilized the dataset in [23], are examined in this section. 13.000 images were trained in three distinct periods, each with a distinct number of images.

The experimental result comprises three separate tasks utilizing image dataset. Task1 contains 6800 images including 3800 images "aca" and 3000 images "n". Task2 consists of 5200 images, with 1400 images "aca" and 3800 images "scc". Task3 consists of 1000 images, with 238 images "aca", 224 images " n", and 210 images "scc" Each task in the proposed model separated the dataset into training and testing sets, Task1 employs 5440 images for training and 1360 for testing, Task2 employs 4160 images for training and 1040 for testing, Task3 employs 800 images for training and 200 for testing. The experimental framework is structured for evaluating performance across varying data compositions and sample sizes, while ensuring balanced train-test ratios in both tasks. The use of diverse category combinations indicates that this may be assessing model efficacy across many classification contexts or examining transfer learning among related yet unique image categories. To tackle a problem of class imbalance, the proposed model generates a training subset that is balanced.

The difference in dataset size (6800, 5200 and 1000) did not affect the results, as balanced subsets were created for each task using subsampling and limited oversampling to ensure equal class representation during training. To achieve balance among the classes, larger classes undergo random subsampling to align with the count of the smallest c lass. Classes with a minimal number of instances specifically that are less than half of the smallest count, experience restricted oversampling, where samples are replicated to enhance their representation. in Task1 balanced training samples is 4836 images for ″aca″ and ″n″. Task2 contains 1120 images of balanced training samples for "aca" and "scc". Task3 contains 627 images of balanced training samples for "aca", " n", and "scc". This leads to a dataset that maintains equilibrium among the classes, thereby averting the model from forming biases in favor of the more prevalent classes.

The proposed model in Table 2 was trained using CNN with fixed hyperparameters (learning rate = 0.0002) and without EWC regularization. Training utilized 6,800 images over 15 epochs, Total time for training is 13 min and 18 s. The average time per epoch is 53 s (fastest is 51 s slowest is 54 s). The training achieved a final loss of 0.0098 with 99.98% accuracy.

The model exhibits high performance and robust generalization ability over the course of 15 epochs. The model exhibited minor fluctuations in performance, recording a test loss of 0.0255 and a test accuracy of 99.85% at epoch 1. However, during epoch 2, there are decreased in accuracy to 99.41%, accompanied by a slight reduction in precision, recall, and F1-score, all getting at 99%. From epoch 5 onward, the model demonstrated exemplary test metrics, achieving 100% for (accuracy, precision, recall, and F1-score), alongside a consistent decline in test loss, which stabilized at 0.0089 from epoch 12 onward. The findings demonstrate that the model achieved rapid convergence while sustaining best classification performance, indicating a high level of learning efficiency and robustness with respect to the test data.

This result in Task1 reaches practical convergence quickly and shows flawless classification on the present test split, with stable training dynamics and tight loss curves strong evidence of effective learning.

After complete Task1, an additional test (Extra Test) was conducted to verify the validity of the training. The classification results displayed in Table 3 demonstrate exemplary model performance on Task 1 for both "aca" and "n". All the chosen 7 images were accurately classified, achieving a 100% accuracy rate for each class and an overall classification accuacy of 100%. This underscores the model's superior discriminative ability and indicates that it has successfully acquired the distinguishing characteristics of each class, with no misclassifications noted.

The confusion matrix illustrated in Figure 2 indicates perfect classification performance of the model on the test set for Task 1, comprising 67 images. All 34 instances of the "aca" class and all 33 instances of the "n" class were accurately identified, with no occurrences of false positives or false negatives. This ideal diagonal matrix, characterized by zero values in the off-diagonal elements, demonstrates

Table 2. Training and testing results for Task1.

Input 6800 image (ACA & N) (5440 image for training & 1360 for testing) (balanced training samples 4836 image)								
Epoch	Train loss	Train accuracy	Test loss	Test accuracy	Test precision	Test recall	Test F1-score	Training time
1	0.3069	96.86%	0.0255	99.85%	100%	100%	100%	51s
2	0.0192	99.77%	0.0215	99.41%	99.00%	99%	99.00%	51s
3	0.0106	100.00%	0.0127	99.93%	100%	100%	100%	52s
4	0.0410	99.05%	0.0316	99.78%	100%	100%	100%	52s
5	0.0263	99.57%	0.0144	100%	100%	100%	100%	53s
6	0.0147	99.88%	0.0103	100%	100%	100%	100%	53s
7	0.0097	100.00%	0.0097	100%	100%	100%	100%	53s
8	0.0100	100.00%	0.0096	100%	100%	100%	100%	53s
9	0.0089	100.00%	0.0090	100%	100%	100%	100%	53s
10	0.0089	100.00%	0.0092	100%	100%	100%	100%	53s
11	0.0090	100.00%	0.0093	100%	100%	100%	100%	53s
12	0.0086	100.00%	0.0089	100%	100%	100%	100%	53s
13	0.0086	100.00%	0.0089	100%	100%	100%	100%	53s
14	0.0087	100.00%	0.0089	100%	100%	100%	100%	54s
15	0.0098	99.98%	0.0089	100%	100%	100%	100%	54s

Table 3. Extra test after complete Task1.

Class	Total images	Correctly classified	Accuracy per class
Aca	34	34	100%
N	33	33	100%
Overall	67	67	100%

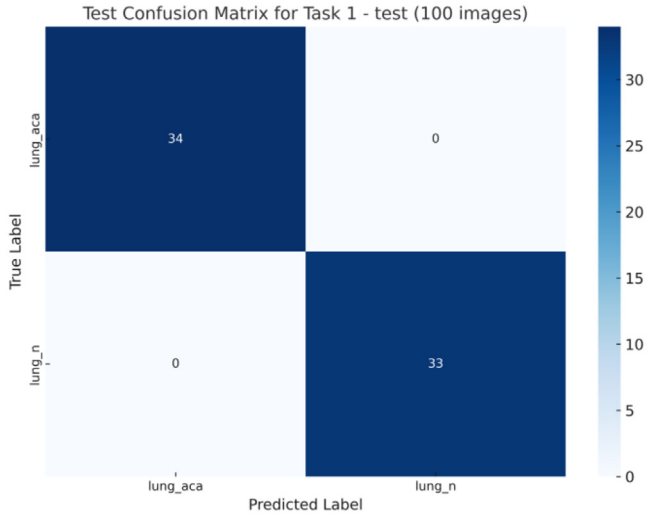

Fig. 2. Confusion matrix for extra test after complete Task1. (Lung_n = normal class "n", lung_aca = adenocarcinoma class "aca").

complete sensitivity and specificity for both classes. The findings visually corroborate the previously reported numerical performance metrics, demonstrating the model's high precision and robust discriminative capability between the two lung tissue categories.

Training on Task2 in Table 4 using EWC with CNN on 5200 histopathology images (4160 for training and 1040 for testing) was completed in 8 min and 42 s, averaging 30 s per epoch (range between 29 and 32 s). The EWC penalty steadily declined from 743 to 515, indicating progressive consolidation of parameters learnt in Task1 while the network acquired new knowledge. Both the EWC λ and the learning rate scheduler were updated rather than held constant:

- Epoch 11: λ dropped from (≈11517 to 11196) and the learning rate was halved to (0.00005) once validation loss plateaued, increasing plasticity just enough to escape a shallow minimum without eroding previously protected weights.
- Epoch 15: a second adjustment (λ ≈ 10498) and learning rate is (0.000025) was triggered when improvement again stalled, enabling finer weight refinements.

Early stopping terminated training at epoch 17 after five consecutive epochs without validation gain (final test loss is 0.0693%, and test accuracy is 96.44%). Together, the gradually declining EWC penalty and dynamically reduced (EWC λ and learning rate) demonstrate a controlled trade-off between stability and plasticity, the proposed model preserved Task1 representations while achieving strong generalization on Task2.

Table 4. Dynamic hyperparameter for Task2.

Input 5200 image (ACA & SCC) (4160 image for training & 1040 for validation) (balanced training samples 1120 image)

Epoch	Train loss	Train accuracy	Test loss	Test accuracy	Test precision	Test recall	Test F1-score	EWC penalty	EWC lambda	Learning rate	Time
1	14864.484	74.6%	0.2973	92.21%	92%	92%	92%	743.193338	14863.866752	0.0001	29s
2	14443.382	89.2%	0.1494	94.9%	95%	95%	95%	722.157739	14443.154785	0.0001	29s
3	14037.938	89.91%	0.1677	90.0%	93%	90%	91%	701.888331	14037.766619	0.0001	29s
4	13645.523	90.89%	0.1263	92.69%	95%	95%	95%	682.268936	13645.378714	0.0001	30s
5	13264.492	90.8%	0.1095	94.71%	95%	94%	95%	663.243091	13264.861816	0.0001	30s
6	12895.49	90.89%	0.1087	93.65%	95%	94%	94%	644.768311	12895.366228	0.0001	31s
7	12536.491	92.28%	0.151	91.06%	94%	94%	94%	626.818920	12536.378409	0.0001	31s
8	12187.493	92.01%	0.0972	93.75%	95%	94%	94%	609.369034	12187.380676	0.0001	30s
9	11848.141	91.62%	0.1574	92.8%	94%	91%	92%	592.401342	11848.026838	0.0001	30s
10	11517.868	92.99%	0.1226	92.02%	94%	92%	94%	575.888551	11517.771014	0.0001	30s
11	11196.358	92.68%	0.108	93.56%	95%	94%	94%	**559.812961**	**11196.259225**	**0.00005**	**30s**
12	10959.762	92.93%	0.1084	96.44%	96%	95%	96%	547.982710	10959.654201	0.00005	30s
13	10804.506	93.12%	0.0942	94.52%	95%	95%	95%	540.222128	10804.424360	0.00005	31s
14	10650.548	93.12%	0.1069	92.88%	94%	93%	93%	532.524924	10650.498483	0.00005	31s
15	10497.975	93.21%	0.1076	95.87%	96%	95%	96%	**524.898781**	**10497.975612**	**0.000025**	**32s**
16	10384.153	93.75%	0.096	95.19%	96%	95%	96%	519.202675	10384.053502	0.000025	31s
17	10308.631	94.73%	0.0724	96.15%	97%	96%	96%	515.428555	10308.571097	0.000025	31s

Table 5. Extra test after complete Task2.

Class	Total images	Correctly classified	Incorrectly classified	Accuracy per class
Aca	472	471	1	99.7%
N	237	237	0	100%
Scc	239	232	7	97%
Total	967	940	Incorrectly classified =8 & Unknown =19	97.2%

The training results displayed distinct and consistent enhancement in model performance over 17 epochs.The training loss steadily decreased from a value of 14,864.48 in epoch 1 to 10,308.63 by epoch 17, indicating effective optimization of the model. Correspondingly, the training accuracy improved substantially, rising from 74.6% to 94.73%, reflecting enhanced model performance over the course of training. The observed trends indicate that the model is acquiring increasingly effective representations over time, demonstrating no indications of overfitting, as reflected by the sustained enhancement across all assessed metrics.

The testing results for Task 2 indicate strong and consistent generalization performance of the model across all 17 epochs. The test loss demonstrates a steady decrease from 0.2973 to 0.0724, reflecting an enhancement in model reliability and assurance. Additionally, the test accuracy shows an enhancement from 92.21% to 96.15%, with comparable advancements noted in test precision, recall, and F1-score, all reaching a maximum of 96% during epoch 17. The metrics indicate a robust and flexible model, exhibiting minimal overfitting while ensuring consistently high and stable performance throughout both training and testing phases. The model demonstrates enhanced classification performance, especially from epoch 12 onwards.

The Extra Test results in Table 5 confirm the strong generalization capability of the proposed model, achieving correct classification for 940 out of 967 images (≈97.2% accuracy). Class-wise precision is particularly high for normal lung tissue (“n”, 100%) and lung adenocarcinoma (“aca”, 99.7%), demonstrating that the decision boundaries for these classes are highly discriminative. In contrast, lung squamous-cell carcinoma (“scc”) shows a slight decline in precision (97%), consistent with the seven misclassifications observed in the confusion matrix.

Detailed error analysis indicates that the misclassified images fall into two categories. Samples predicted with confidence above 0.5 but assigned to the wrong class indicate partial feature recognition without full certainty. By contrast, samples with confidence below 0.5 were

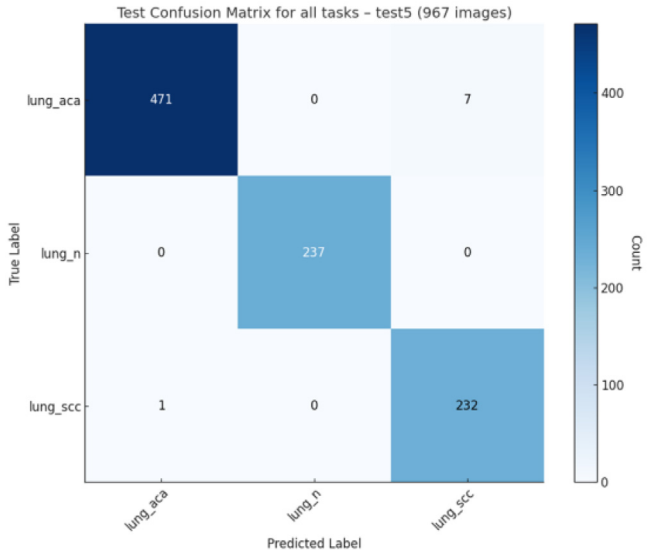

Fig. 3. Confusion matrix for extra test after complete Task2.

assigned to the ″unknown class,″ preventing unreliable predictions from being forced into predefined categories. In total, eight samples were misclassified, while 19 were routed to the unknown group, ensuring safer handling of ambiguous or out-of-distribution inputs. The ″unknown class″ is not part of the training phase; rather, it is introduced during testing as a mechanism to isolate samples that the model cannot confidently classify. In practice, the unknown class prevents forcing uncertain predictions into incorrect categories and ensures that such ambiguous or out-of-distribution inputs are grouped separately. This property is critical for real-world diagnostic pipelines, where models must be able to withhold uncertain decisions rather than risking erroneous clinical outputs.

Although the overall accuracy balance performance and memory protection of our model (97.2%) is slightly lower than some reported results (98–99.6%), this small reduction is acceptable in the context of continual learning. Unlike conventional single-task training, our approach is explicitly designed to preserve knowledge across sequential tasks and mitigate catastrophic forgetting. The marginal drop in accuracy is therefore the trade-off for achieving higher stability, adaptability to new tasks, and long-term retention of prior knowledge. In practical applications, this balance between accuracy and continual learning capability is often more valuable than maximizing accuracy on a single isolated task.

The confusion matrix in Figure 3 illustrates the model performance on an extra test set comprising 967 images across three lung pathology classes. The model demonstrates excellent classification accuracy, with most predictions aligning perfectly along the diagonal, indicating correct classifications. Classes ″n″ and ″aca″ indicate a high level of identification accuracy, with only seven misclassified of ″scc″ images and one misclassified of ″aca″ image. This indicates a significant level of class separability and the efficient acquisition of discriminative features. Significantly, there are no false positives detected in the "n" class, which additional underscores the model's robustness. Limiting off diagonal entries further validates the minimal class confusion, reinforcing the model dependability in clinical or diagnostic contexts.

Upon completing the initial two tasks, new data emerged, necessitating a third round of training. Table 6 presents a summary of the dynamic behavior of the training and evaluation metrics over 12 epochs for Task 3, including the progression of the EWC-related parameters and the learning rate.

The results indicate a significant convergence, with the training loss decreasing from 0.34 to 0.03, while the training accuracy rose from 89% to almost 99%. The test accuracy stabilized at approximately 99% from the middle epochs onward, with precision, recall, and F1-score achieving nearly perfect values, demonstrating robust generalization. The EWC penalty exhibited a consistent decline from 743 in the initial epoch to approximately 515 in the concluding epoch, while the λ parameter decreased from roughly 14,863 to nearly 10,300. This adaptive adjustment demonstrates the equilibrium between safeguarding established knowledge and facilitating flexibility for the new task. The learning rate was maintained at a constant value of 0.0002 for most of the epochs, before being adjusted to 0.0001 during the final stage for fine-tuning purposes. The results collectively demonstrate that the model effectively incorporated new information while ensuring stability across sequential tasks an important characteristic for clinical diagnostic pipelines, where the ability to adapt to new data and retain prior knowledge is vital for dependable decision-making. The results in Table 7 and Figure 4 show that the model attains an overall accuracy of 98.96% (665/672) on the extra test set. Only seven images were not assigned to the correct class: four were misclassified into incorrect labels while three were routed to the unknown class. The ″unknown class″ introduced at test time rather than during training captures low confidence or out-of-distribution samples and thus prevents unreliable predictions from being forced into predefined categories. This conservative handling of ambiguous cases enhances diagnostic safety by allowing the system to withhold uncertain decisions instead of producing potentially misleading clinical labels.

The results in Table 7 and Figure 4 show that the model attains an overall accuracy of 98.96% (665/672) on the extra test set. Only seven images were not assigned to the correct class: four were misclassified into incorrect labels while three were routed to the unknown class. The ″unknown class″ introduced at test time rather than during training captures low confidence or out-of-distribution samples and thus prevents unreliable predictions from being forced into predefined categories. This conservative handling of ambiguous cases enhances diagnostic safety by allowing the system to withhold uncertain decisions instead of producing potentially misleading clinical labels.

Figure 5 presents the normalized distribution of parameter importance across various lung tissue classes (″aca″, ″n″, and ″scc″) for the three tasks. The findings indicate clear patterns of parameter activation based on

Table 6. Dynamic hyperparameter for Task3.

Input 1000 image (ACA &N& SCC) (800 image for training & 200 for validation) (balanced training samples 627 image)											
Epoch	Train loss	Train accuracy	Test loss	Test accuracy	Test precision	Test recall	Test F1-score	EWC penalty	EWC lambda	Learning rate	Time
1	0.3425	89.31%	0.0808	97.5%	98%	97%	98%	743.19	14863.87	0.0002	9s
2	0.2153	93.3%	0.0906	96.5%	97%	96%	96%	701.05	14021.04	0.0002	9s
3	0.1177	94.74%	0.0759	98%	98%	98%	98%	665.42	13308.4	0.0002	9s
4	0.0906	96.17%	0.0669	98%	98%	98%	98%	642.33	12846.6	0.0002	9s
5	0.0695	96.97%	0.0654	98%	98%	98%	98%	620.17	12403.4	0.0002	9s
6	0.0554	97.77%	0.0578	98%	98%	98%	98%	597.41	11948.2	0.0002	9s
7	0.0605	97.13%	0.041	99%	99%	99%	99%	575.11	11502.2	0.0002	9s
8	0.0403	98.56%	0.0499	98%	99%	97%	99%	553.22	11064.4	0.0002	9s
9	0.0631	97.13%	0.0645	97.5%	98%	97%	98%	536.88	10737.6	0.0002	9s
10	0.0724	97.77%	0.0483	98.5%	99%	98%	99%	525.44	10508.8	0.0001	9s
11	0.0634	98.56%	0.0432	98.5%	99%	98%	99%	520.22	10404.4	0.0001	9s
12	0.0356	98.88%	0.0443	99%	99%	99%	99%	515.43	10308.5	0.0001	9s

Table 7. Extra test after complete Task3.

Class	Total images	Correctly classified	Incorrectly classified	Accuracy per class
Aca	238	234	1	98.3%
N	224	224	0	100%
Scc	210	207	3	98.5%
Total	672	665	Incorrectly classified =4 & Unknown =3	98.96%

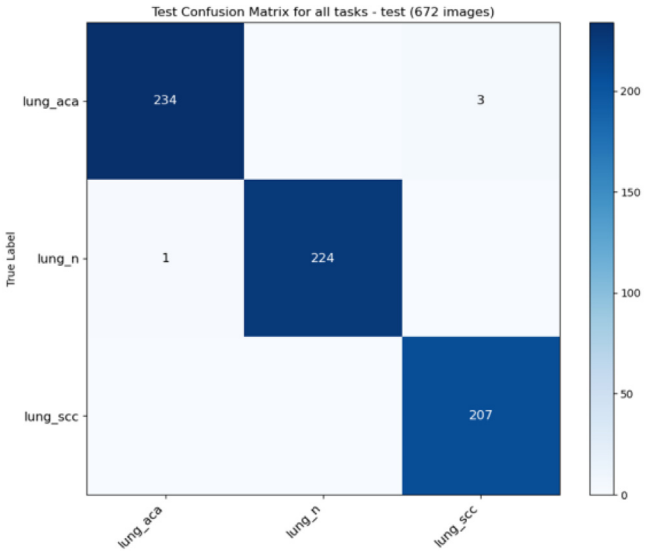

Fig. 4. Confusion matrix for extra test after Task3.

the classification objective. In Task 1, parameters such as bn1.weight and conv1.weight play a crucial role in differentiating ″aca″ from ″n″, indicating that the model predominantly depends on initial convolutional and normalization layers to distinguish tumor tissue from normal tissue. In Task 2, the emphasis is placed on more intricate parameters such as bn2.weight and bn2.bias for distinguishing ″aca″ from ″scc″, suggesting that the model identifies finer distinctions between cancer subtypes in the deeper layers. In Task 3, the significance is more evenly spread across various parameters for ″aca″ and ″n″, whereas ″scc″ demonstrates a notable predominance in specific parameters, highlighting the model's capacity to effectively identify its uniquely characteristic features. The findings illustrate the model's adaptability in its reliance on various layers based on the classification task, emphasizing the importance of hierarchical feature representations in differentiating between normal tissue and cancer subtypes.

Table 8 provides a comparative analysis of the proposed model in relation to previously reported approaches. The ResNet50-EWC model demonstrated an impressive accuracy of 98.96%, surpassing or matching the performance of various previous methods, including the Custom 9-layer CNN at 97.20%, Shallow 3-layer CNN at 97.89%, Residual VGG-19 + CLAHE at 97.73%, and Baseline CNN at 95.00%. Although techniques like Ensemble CNN with Grad-CAM visualization (up to 99.69%) and CNN-EWC (up to 99.6%) have demonstrated marginally higher peak accuracies, the proposed model presents distinct advantages. This approach integrates EWC with ResNet50, facilitating continual learning and dynamic adaptation without the need for retraining from the beginning. The characteristics improve the model's generalizability, mitigate the risk of

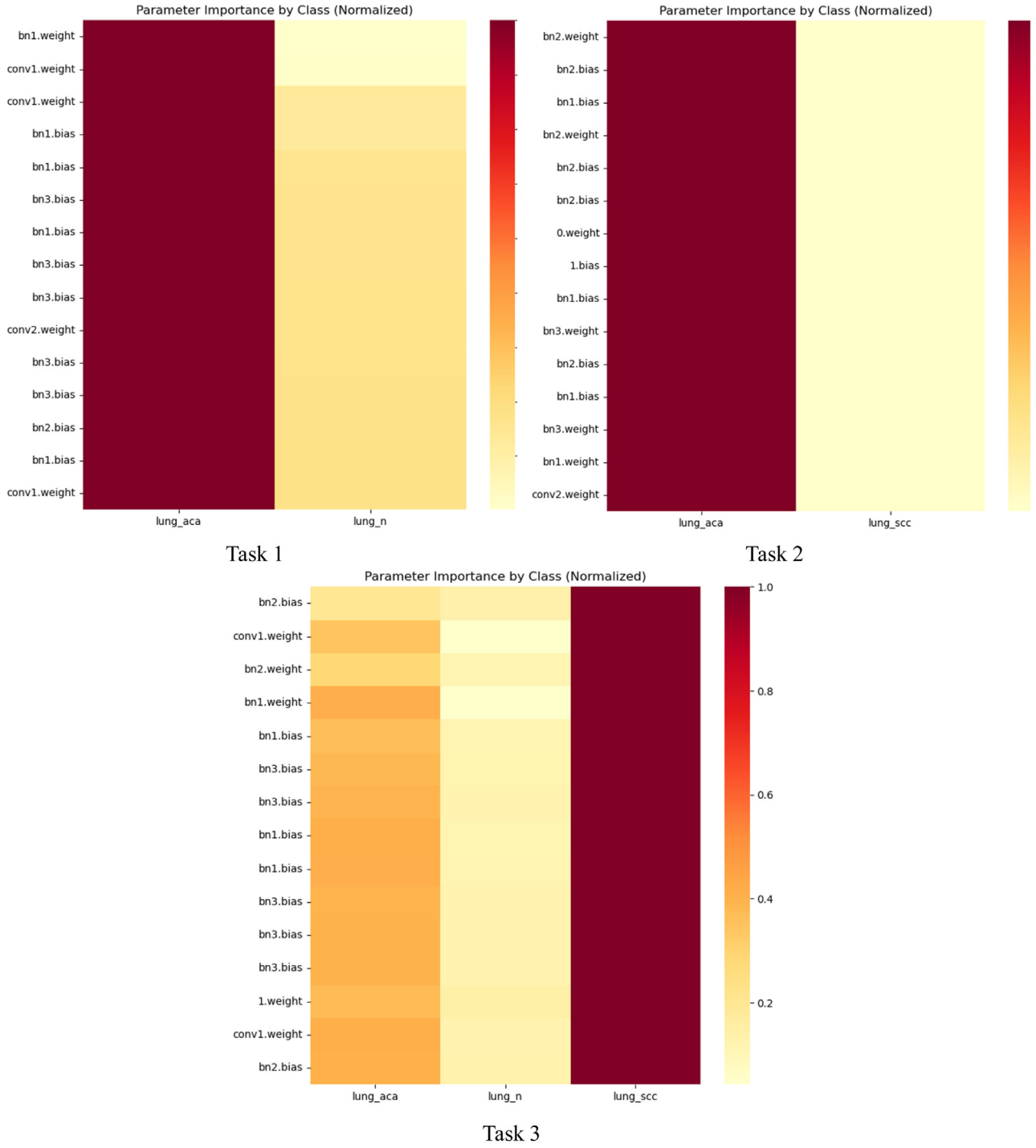

Fig. 5. Normalized parameter importance across classes for three tasks.

overfitting, and enhance training efficiency, thereby increasing its suitability for real-world clinical applications where data distributions change over time. Moreover, by integrating EWC with ResNet50, the model not only sustains high accuracy but also effectively preserves previously acquired knowledge. This balance between performance and memory protection ensures reliable adaptation to new data while minimizing catastrophic forgetting.

5 Limitations

The proposed model exhibits certain limitations. The accuracy, while competitive, falls marginally short of the leading ensemble and transfer learning methodologies. The training is confined to a specific selection of lung cancer subtypes, potentially diminishing its applicability to more extensive histopathology datasets. Furthermore, the

Table 8. Comparison with previous research.

No.	References	Method Used	Accuracy Achieved
1	[14]	Custom 9-layer CNN	97.20 %
2	[15]	Shallow 3-layer CNN	97.89 %
3	[16]	Ensemble CNN + Grad-CAM visualization	97.11–99.69 %
4	[17]	Residual VGG-19 + CLAHE	97.73 %
5	[18]	Baseline CNN	95.00 %
6	[19]	Transfer-learned ResNet-50 CNN	98.00 %
7	[20]	CNN-EWC	98– 99.6 %
8	Proposed Model	Resnet50-EWC	98.96 %

adaptive hyperparameter scheduler, although it minimizes the need for manual tuning, may not consistently attain best configurations when dealing with highly heterogeneous or imbalanced data.

6 Conclusions

This research introduces an architecture grounded in ResNet50, augmented by EWC and an adaptive hyperparameter scheduler, aimed at the classification of lung cancer histopathology. The proposed model attains an impressive accuracy of 98.96% in Task3, accompanied by three significant contributions: (i) resilience against catastrophic forgetting when new tissue subclasses are incorporated, (ii) automated hyperparameter tuning that minimizes the need for manual experimentation, and (iii) a reduced computational footprint relative to ensemble methods, facilitating quicker and more resource-efficient deployment. The benefits contribute to broader applicability, mitigate the risk of overfitting, and facilitate the integration of practical clinical workflows amidst changing data distributions.

Nonetheless, certain constraints persist. The model demonstrates competitive accuracy, albeit marginally trailing behind the leading ensemble and transfer learning methodologies. The training is confined to a specific subset of lung cancer subtypes, which may limit its relevance to more extensive histopathology datasets. Furthermore, although the adaptive scheduler mitigates the need for manual tuning, significant heterogeneity within datasets may continue to pose challenges to achieving best performance.

Future endeavors may concentrate on: (i) refining model precision via architectural advancements such as attention mechanisms or multi-scale feature extraction, (ii) broadening the model to encompass additional lung cancer subtypes and associated tissue categories for enhanced clinical significance, and (iii) incorporating semi-supervised or self-supervised continual learning methodologies to utilize unlabeled data and further augment resilience to dynamic datasets.

Funding

The authors received no specific funding for this study.

Conflicts of interest

The authors declare that they have no conflicts of interest.

Data availability statement

The dataset used in this study is publicly available. The dataset can be accessed from [23].

Author contribution statement

Zainab Muhammed Jameel: Implementation of codes, analysis of results and comparison with previous works. Belal Al-Khateeb: Design of the method and review of the paper.

Ethics approval

This study does not involve humans, animals, or sensitive data that require ethical approval.

References

1. R.L. Siegel, A.N. Giaquinto, A. Jemal, Cancer statistics, 2024, CA: Can. J. Clin. **74**, 12–49 (2024)
2. W.D. Travis, E. Brambilla, A.P. Burke, A. Marx, A.G. Nicholson, Introduction to the 2015 World Health Organization classification of tumors of the lung, pleura, thymus, and heart, J. Thorac. Oncol. **10**, 1240–1242 (2015)
3. A. Homeyer et al., Recommendations on compiling test datasets for evaluating artificial intelligence solutions in pathology, Mod. Pathol. **35**, 1759–1769 (2022)
4. G. Litjens et al., A survey on deep learning in medical image analysis, Med. Image Anal. **42**, 60–88 (2017)
5. H.Y. Chang et al., Artificial intelligence in pathology, J. Pathol. Trans. Med. **53**, 1–12 (2019)
6. D.C. Ciresan, A. Giusti, L.M. Gambardella, J. Schmidhuber, Mitosis detection in breast cancer histology images with deep neural networks, Medical Image Computing and Computer-Assisted Intervention – MICCAI **2013**, 411–418 (2013)
7. B. Ehteshami Bejnordi et al., Diagnostic assessment of deep learning algorithms for detection of lymph node metastases in women with breast cancer, JAMA **318**, 2199 (2017)

8. C.-M. Hsu, C.-C. Hsu, Z.-M. Hsu, F.-Y. Shih, M.-L. Chang, T.-H. Chen, Colorectal polyp image detection and classification through grayscale images and deep learning, Sensors **21**, 5995 (2021)
9. S. Mangal, A. Chaurasia, A. Khajanchi, Convolution neural networks for diagnosing colon and lung cancer histopathological images, arXiv.org, Sep. 08, 2020. https://arxiv.org/abs/2009.03878
10. J. Civit-Masot, A. Bañuls-Beaterio, M. Domínguez-Morales, M. Rivas-Pérez, L. Muñoz-Saavedra, J.M. Rodríguez Corral, Non-small cell lung cancer diagnosis aid with histopathological images using explainable deep learning techniques, Comput. Methods Programs Biomed. **226**, 107108 (2022)
11. L. Song, H. Wang, Z.J. Wang, Decoupling multi-task causality for improved skin lesion segmentation and classification, Pattern Recognit. **133**, 108995 (2023)
12. J. Kirkpatrick et al., Overcoming catastrophic forgetting in neural networks, Proc. Natl. Acad. Sci. **114**, 3521–3526 (2017)
13. J. Chen, Comparative analysis of machine learning models for lung cancer prediction, in: *2023 IEEE International Conference on Image Processing and Computer Applications (ICIPCA)*, 2023, pp. 242–246
14. B.K. Hatuwal, H.C. Thapa, Lung cancer detection using convolutional neural network on histopathological images, Int. J. Comput. Trends Technol. **68**, 21–24 (2020)
15. S. Mangal, A. Chaurasia, A. Khajanchi, Convolution neural networks for diagnosing colon and lung cancer histopathological images, arXiv.org, Sep. 08, 2020. https://arxiv.org/abs/2009.03878
16. J. Civit-Masot, A. Bañuls-Beaterio, M. Domínguez-Morales, M. Rivas-Pérez, L. Muñoz-Saavedra, J.M. Rodríguez Corral, Non-small cell lung cancer diagnosis aid with histopathological images using explainable deep learning techniques, Comput. Methods Programs Biomed. **226**, 107108 (2022)
17. S. Wadekar, D.K. Singh, A modified convolutional neural network framework for categorizing lung cell histopathological image based on residual network, Healthcare Analytics, **4**, 100224–100224 (2023)
18. R.D. Mohalder, K.A. Hossain, J.P. Sarkar, L. Paul, M. Raihan, K.H. Talukder, Lung cancer detection from histopathological images using deep learning, in: *Machine Intelligence and Emerging Technologies, Lecture Notes of the Institute for Computer Sciences, Social Informatics and Telecommunications Engineering*, vol. 490, 2023, pp. 201–212
19. M.F. Khatun, M.R. Ajmain, Md. Assaduzzaman, A deep learning approach to detect and classification of lung cancer, Int. Conf. Adv. Technol. (ICONAT), 1–6 (2023)
20. Z. Muhammed, B. Al-Khateeb, CNN-EWC: a continuous deep learning approach for lung cancer classification, J. Intell. Syst. **34**, 20240541 (2025)
21. Y. He, J. Henze, B. Sick, Continuous learning of deep neural networks to improve forecasts for regional energy markets, IFAC-PapersOnLine, **53**, 12175–12182 (2020)
22. M. Lenga, H. Schulz, A. Saalbach, Continual learning for domain adaptation in chest X-ray classification, arXiv.org, 2020. https://arxiv.org/abs/2001.05922
23. Lung and Colon Cancer Histopathological Images, www.kaggle.com. https://www.kaggle.com/datasets/andrewmvd/lung-and-colon-cancer-histopathological-images

Cite this article as: Zainab Muhammed, Belal Al-Khateeb, Adaptive hyperparameter and regularization strategies for improved elastic weight consolidation in lung cancer classification. Int. J. Simul. Multidisci. Des. Optim. **17**, 15 (2026). https://doi.org/10.1051/smdo/2026014