Multi-Approach Deep Learning for Histopathologic Cancer Detection: Supervised, Unsupervised, and State-of-the-Art Models

Damien Liscio
Centennial College

Fan Yang
Centennial College

Harpreet Singh Dhanda
Centennial College

Sophia Ojegba
Centennial College

***Abstract*— Accurate and early detection of metastatic cancer is critical in medical diagnosis and treatment planning. In this study, we explore deep learning approaches to automate the detection of cancerous tissues using histopathologic images. We use the publicly available Histopathologic Cancer Detection dataset from Kaggle to evaluate three models: a custom Convolutional Neural Network (CNN) for supervised learning, a Generative Adversarial Network (GAN) for unsupervised data augmentation, and a fine-tuned EfficientNet model representing state-of-the-art performance. Our results demonstrate that data augmentation using GANs significantly improves the robustness of the CNN, while EfficientNet achieves the best performance across all metrics, validating the power of transfer learning for complex biomedical image classification.**

***Keywords*— cancer detection, convolutional neural network, supervised learning, deep learning, histopathologic image classification**

I. Introduction

Histopathologic cancer detection is a critical application of medical image classification, with direct implications for diagnostic accuracy and treatment planning. In this project, we explore multiple deep learning paradigms to address the binary classification of microscopic tissue images into cancerous or non-cancerous classes.

The dataset used includes over 220,000 .tif images and a CSV file containing binary labels. The objective is to develop models capable of learning discriminative patterns from this high-volume, high-resolution image data. To approach this task comprehensively, our group implemented and evaluated three experiments:

1. A supervised learning model designed and trained from scratch using a custom Convolutional Neural Network (CNN),
2. An unsupervised learning model leveraging Generative Adversarial Networks (GANs) for data augmentation,
3. A state-of-the-art model based on the pre-trained EfficientNet architecture using transfer learning

This paper presents the methodology and results of the supervised learning, unsupervised learning and state-of-the-art approach.

II. Methodology

A. Supervised Learning

For the supervised learning experiment, we implemented a convolutional neural network (CNN) from scratch using the PyTorch framework. The goal was to perform binary classification of histopathologic image patches by predicting the presence (1) or absence (0) of cancerous cells.

The dataset consists of over 220,000 .tif image patches and a corresponding label file in CSV format. Each image was resized to 64×64 pixels and normalized to improve convergence. A stratified split was used to divide the data into 80% training and 20% validation sets with seed = 42.

The CNN architecture included three convolutional blocks, each with a convolutional layer, batch normalization, ReLU activation, and max pooling. Two architectural variations were evaluated: one with filter sizes (32, 64, 128) and another with (64, 128, 256). After the convolutional layers, the feature maps were flattened and passed through a fully connected layer with 256 neurons and a dropout layer for regularization. The final output layer consisted of a single neuron activated by a sigmoid function.

To assess the influence of hyperparameters, we tested combinations of dropout rates (0.3, 0.5) and learning rates (0.001, 0.0005). The model was trained using the Adam optimizer and the binary cross-entropy loss with logits (BCEWithLogitsLoss). Each configuration was trained for two epochs to balance time constraints with experimentation needs. Validation accuracy was used as the primary metric, and the best-performing model was saved and used to generate predictions on the test set.

All experiment configurations, results, and outputs were saved in a structured results/ directory, including a log of validation accuracies and sample visualization of validation images..

B. Unsupervised Learning

In the unsupervised learning experiment, we developed a Generative Adversarial Network (GAN) to synthetically generate histopathologic images that resemble real cancerous and non-cancerous tissues. The primary objectives of this approach were to explore the generative capacity of deep neural networks in the absence of label supervision.

We used a balanced subset of the dataset consisting of 20,000 grayscale images (10,000 per class), resized to 96×96 pixels, and normalized pixel intensities to a [-1, 1] range. This preprocessing ensured consistency and improved convergence during training. The GAN was implemented using TensorFlow and Keras, with careful architectural and training configurations to stabilize the notoriously challenging adversarial optimization.

The Generator network took a 128-dimensional noise vector as input and passed it through a sequence of fully connected, batch normalization, LeakyReLU, and transposed convolution layers to upsample into a 96×96×1 grayscale image. Conversely, the Discriminator was a convolutional network designed to distinguish between real and fake images, consisting of multiple Conv2D layers with LeakyReLU activation and spatial dropout for regularization, followed by a final sigmoid classifier.

To mitigate training instability, several strategies were employed: label smoothing (assigning real images labels in the range [0.8, 0.9] and fake images [0.0, 0.1]), input noise injection (Gaussian noise added to both real and generated images), and gradient clipping. The GAN was trained for 50 epochs using the Adam optimizer with an exponentially decaying learning rate and β_1 set to 0.5. Both Generator and Discriminator losses were monitored across epochs.

Images were generated and saved at regular intervals to assess progression.

C. State-of-the-Art Models

We used EfficientNetB0 as the state-of-the-art model in two setups: (1) trained from scratch and (2) with transfer learning using ImageNet weights. Images were resized to 96×96, normalized, and converted to RGB.

A reduced dataset of 20,000 balanced images was created. Data augmentation included random flips, rotations, zooms, translations, and contrast adjustments.

The model had a custom head with dense layers, batch normalization, and dropout. For transfer learning, the EfficientNet base was first frozen, then partially unfrozen for fine-tuning. Training used the Adam optimizer, early stopping, and was evaluated using accuracy, loss, and AUC.

Transfer learning outperformed training from scratch, confirming its effectiveness on medical image data with limited labels.

III. Result

A. Supervised Learning

Two variants of a convolutional neural network were trained on the complete labeled dataset. The first configuration used three convolutional layers with filter sizes of 32, 64, and 128, a dropout rate of 0.5, and a learning rate of 0.001. This model demonstrated efficient convergence, with an average training loss of 0.3928 in the first epoch and 0.3140 in the second. It achieved a validation accuracy of 89.7% and was selected as the best-performing model.

The second configuration employed deeper convolutional layers with filter sizes of 64, 128, and 256, a reduced dropout rate of 0.3, and a learning rate of 0.0005. While this model also showed a consistent reduction in training loss across epochs (0.3772 in the first and 0.2963 in the second), it achieved a lower validation accuracy of 80.1%. Moreover, it required substantially more training time, averaging 4.1 images per second, indicating that the added depth did not yield better performance under the given hyperparameters.

The model from the first configuration was selected and used to generate predictions on the unseen test dataset. The results were exported to results/submission.csv, and a sample visualization of correctly predicted validation images was saved as sample_images.png to support qualitative analysis.

B. Unsupervised Learning

When We trained a Deep Convolutional GAN (DCGAN) for 50 epochs with a batch size of 64 and a latent dimension of 128. Both the generator and discriminator were optimized using the Adam optimizer with an exponentially decaying learning rate starting at 0.0002 and $\beta_1 = 0.5$. The loss function used was Binary Crossentropy.

Training Dynamics:

Generator Loss decreased sharply in the initial epochs, stabilizing around 0.76–0.78, indicating improved capability to generate realistic samples.

Discriminator Loss increased early and then plateaued near 1.37–1.38, suggesting it reached a steady state, struggling to distinguish between real and fake inputs due to improved generator quality.

The training stabilized after epoch 10, with both networks reaching equilibrium, showing the GAN avoided mode collapse and maintained adversarial balance.

Image Quality & Observations:

Generated images gradually improved in realism and structure. Initial outputs were blurry or noisy, but later epochs produced images with clearer tissue textures and patterns.

The final generator demonstrated good feature understanding, effectively synthesizing 96×96 grayscale histopathologic images similar in visual characteristics to the real dataset.performance.

C. State-of-the-Art Models

The evaluation of the EfficientNet model was conducted through both training from scratch and using transfer learning.

When trained from scratch, the model showed a loss of 17.5424, an accuracy of 51.03%, and an AUC score of 0.4594. These results suggest that the model struggled to learn the underlying patterns effectively from the initial training, with a relatively low accuracy and AUC score indicating limited performance.

In contrast, using transfer learning, the model's performance improved significantly. The loss decreased to 0.6800, and the accuracy increased to 58.70%, with an AUC of 0.6248. The use of pre-trained weights allowed the model to learn more efficiently, enhancing both its generalization ability and classification performance.

IV. Contribution

This project represents a collective effort from four contributors, each bringing valuable expertise to the development and evaluation of the deep learning models for histopathologic cancer detection.

Damien Liscio: Led the design and implementation of the custom Convolutional Neural Network (CNN) for the supervised learning task. This included the preprocessing of the dataset, defining the CNN architecture, and evaluating the impact of various hyperparameters on the model's performance.

Fan Yang: Focused on the unsupervised learning approach, developing the Generative Adversarial Network (GAN) to generate synthetic histopathologic images. They were responsible for configuring the GAN architecture, handling training dynamics, and ensuring stability throughout the process.

Sophia Ojegba: Managed the state-of-the-art model, EfficientNet, including both the training from scratch and transfer learning experiments. They fine-tuned the pre-trained model, conducted data augmentation, and assessed the performance improvements with transfer learning.

Harpreet Singh Dhanda: Contributed to the overall evaluation, analysis, and presentation of results. This included comparing the performances of all three models, summarizing key findings, and contributing to the writing of the methodology and results sections.

Each member contributed to the data preparation, coding, model evaluation, and writing of the paper, making this a collaborative effort to push forward the understanding of deep learning applications in cancer detection.

V. Reference

[1] M. Tan and Q. Le, "EfficientNet: Rethinking Model Scaling for Convolutional Neural Networks," *ICML*, 2019.
[2] I. Goodfellow et al., "Generative Adversarial Nets," *NIPS*, 2014.
[3] Kaggle, "Histopathologic Cancer Detection," [Online]. Available: https://www.kaggle.com/competitions/histopathologic-cancer-detection/data