

## Final Report

### A. Abstract

This study delves into the realm of transfer learning within deep learning (DL) to address the issue of prolonged training time. Two convolutional neural network models are trained on a medical-related dataset, one trained from scratch with randomly initialized weights and another with pre-trained ImageNet weights. Moreover, transfer learning is utilized to extract features from diverse datasets by pre-trained DL models and then Machine Learning classifiers are employed for feature classification. The study explores how transfer learning success depends on the similarity between the pre-trained and target datasets. Results demonstrate that employing ImageNet weights for training yields superior performance, with faster convergence. Furthermore, when comparing pre-trained models, one trained on a limited-size medical dataset and the other on a large, diverse image dataset, the latter excels at classifying new diverse datasets, while the specialized pre-trained model excels only on similar dataset to its pre-trained dataset.

### B. Introduction

Deep learning (DL), using convolutional neural networks (CNNs), has been effectively applied to computer vision problems such as image classification, segmentation, and object detection [1]. CNNs excel at feature extraction, offering advantages such as dimensionality reduction and facilitating transfer learning [2]. Transfer learning (TL) involves utilizing pre-learned universal features from large datasets and fine-tuning pre-trained models based on specific datasets [3]. However, computational complexity, domain-specific issues, real-time constraints, and generalization deficits are ongoing challenges. Promising solutions include Model Compression [4], Domain Adaptation [5,6], Meta-Learning [7], and Regularization Techniques. Methods like pruning and quantization reduce neural network size and computational needs [8–10]. Knowledge distillation trains a smaller 'student' network to mimic a larger pre-trained 'teacher' network [11–13]. Meanwhile, Meta-learning allows a model to adapt to new tasks using limited-size data by leveraging knowledge from similar tasks. The project's objectives include assessing model selection and initialization methods' impact on classification, evaluating feature extraction efficiency in pre-trained models, and assessing their adaptability across domains. [14–17]. In addition, machine learning (ML) models could classify extracted features [18], joining DL's feature representation with ML's classification advantages. The project involves implementing two CNN models on the Colorectal Cancer (CRC) dataset, one with random weight initialization (denoted as CRC-Enc) and another using TL with ImageNet weights (denoted as Img-Enc). The Img-Enc model ex-

hibited superior classification accuracy, benefiting from its prior knowledge of image features, and fine-tuning Img-Enc with the CRC dataset further enhanced its performance. These findings underscored the value of transfer learning in expediting convergence and improving classification outcomes for specialized tasks. In the subsequent task, both CRC-Enc and a ResNet18 model pre-trained on ImageNet (Pre-Img) were employed for feature extraction and classification on prostate cancer and animal face datasets Support Vector Machine (SVM), K-Nearest Neighbors (KNN), and Random Forest (RF). Various evaluation metrics are employed to assess classification performance, encompassing measures such as accuracy, precision, recall, and the F1 score. These metrics collectively provide a comprehensive view of the models' effectiveness in correctly classifying instances, balancing considerations of both true positives and false positives, thereby offering a robust evaluation of the classification outcomes. The results show that both CRC-Enc and Pre-Img models perform well in classifying the prostate cancer dataset. Pre-Img's slightly higher accuracy with SVM and KNN suggests that the broader knowledge gained from the extensive ImageNet dataset benefits the classification of the smaller specialized prostate cancer dataset. Conversely, CRC-Enc features were less effective in animal face classification but Pre-Img easily managed to classify them. Additionally, SVM consistently outperforming KNN and RF, resulting in an average improvement of approximately 30% in domain-specific tasks. These results underscore the efficacy of leveraging pre-trained models and the pivotal role of model choice in optimizing performance across diverse domains, contributing valuable insights to the field of machine learning and image classification.

Various studies in the literature employ transfer learning-based DL models for various medical image analysis tasks [19–26]. In [19] and [20], deep learning models based on transfer learning techniques are developed for breast cancer detection and diagnosis. Both studies utilize pre-trained ResNet-50 architectures on ImageNet to extract features from breast cancer datasets. Evaluation metrics such as accuracy, sensitivity, specificity, precision, and F-score are employed to assess the performance of these models. The study in [21] demonstrates the detection of pancreatic cancer in hematoxylin and eosin (H&E) sections using convolutional neural networks and deep transfer learning. The ResNet18 network achieved a test accuracy of approximately 94% for five-class classification. The paper in [22] introduces the combination of deep neural network model compression methods to enhance compression results. These methods were tested on the ResNet18 model pretrained on the NCT-CRC-HE-100K dataset. In [23], the study investigates deep transfer learning for lung cancer

000  
001  
002  
003  
004  
005  
006  
007  
008  
009  
010  
011  
012  
013  
014  
015  
016  
017  
018  
019  
020  
021  
022  
023  
024  
025  
026  
027  
028  
029  
030  
031  
032  
033  
034  
035  
036  
037  
038  
039  
040  
041  
042  
043  
044  
045  
046  
047  
048  
049  
050  
051  
052  
053  
054  
055  
056  
057  
058  
059  
060  
061  
062  
063  
064  
065  
066  
067  
068  
069  
070  
071  
072  
073  
074  
075  
076  
077  
078  
079  
080  
081  
082  
083  
084  
085  
086  
087  
088  
089  
090  
091  
092  
093  
094  
095  
096  
097  
098  
099  
100  
101  
102  
103  
104  
105  
106  
107

108  
109  
110  
111  
112  
113  
114  
115  
116  
117  
118  
119  
120  
121  
122  
123  
124  
125  
126  
127  
128  
129  
130  
131  
132  
133  
134  
135  
136  
137  
138  
139  
140  
141  
142  
143  
144  
145  
146  
147  
148  
149  
150  
151  
152  
153  
154  
155  
156  
157  
158  
159  
160  
161  
classification using ResNet50 as a feature extractor on the Lung Image Database. Deep features are then classified using various classifiers, including Naive Bayes, MultiLayer Perceptron, Support Vector Machine, k-Nearest Neighbors, and Random Forest. The best performance is achieved with SVM, attaining an accuracy of 88.41%. The paper in [25] investigates the benefits of using domain-specific pretrained weights for deep learning models when classifying histopathological tissue images. Their research reveals that domain-specific pretrained weights outperform ImageNet pretrained weights, yielding higher performance and shorter training times on histopathology datasets.

## C. Methodology

There are three datasets in this project. The first dataset, **NCT-CRC-HE-100K** [27], contains 6000 images from 100k colorectal cancer (CRC) images, taken from hematoxylin & eosin (H&E) stained histological images of human colorectal cancer and normal tissues. The images were manually extracted from formalin-fixed paraffin-embedded (FFPE) samples by the National Tumor Diseases Center. The classes of tissues are Smooth Muscle, Normal Colon Mucosa, and Cancer-Associated Stroma. The second dataset has 6000 images extracted from 120k patches of the **prostate cancer** dataset [28]. Patches were gathered from six subsets from four distinct institutions, each digitized using different scanner systems. The types of tissues are Prostate Cancer Tumor, Benign Glandular Prostate, and Benign Non-Glandular Prostate. The third dataset, extracted from 16k **animal faces** images, comprises 6000 images showcasing diverse cats, dogs and wildlife [29]. The images were collected with permissive licenses from the Flickr1 and Pixabay2 websites. The statistics of the datasets are listed in Table 1. Before training, the datasets undergo essential preprocessing steps aimed at strengthening the model's robustness. These steps include normalization, which not only standardizes the pixel values to a consistent range but also helps in reducing sensitivity to variations in lighting and color. Data augmentation techniques such as random flipping and rotation are also applied to diversify the training dataset. Furthermore, all images are resized to uniform dimensions of 224x224 pixels. These preprocessing measures collectively enhance the model's performance and enable it to better handle variations in the input data. For instance, the CRC dataset, displayed before and after preprocessing in Fig. 1. The datasets are found to have a balanced class distribution which is evident by displaying the number of images for each Fig 3. All datasets will undergo division into training, testing, and validation sets using an 80-10-10 split ratio.

The project consists of two primary tasks. The first task is about implementing two CNN-based models using the CRC dataset. The Residual Network architecture, specific-

Datasets	Colorectal Cancer		Prostate Cancer		Animals	
	All	Subset	All	Subset	All	Subset
Image size	224x224x3		300x300x3		512x512x3	
Image Format	TIFF		JPG		JPG	
Image count	100,000	6,000	120,000	6,000	16,000	6,000
Class count	8	3	3	3	3	3

Table 1. Datasets Statistics &amp; Attributes

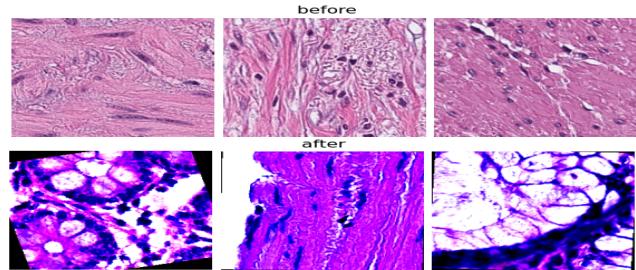


Figure 1. Task 1 dataset before and after preprocessing.

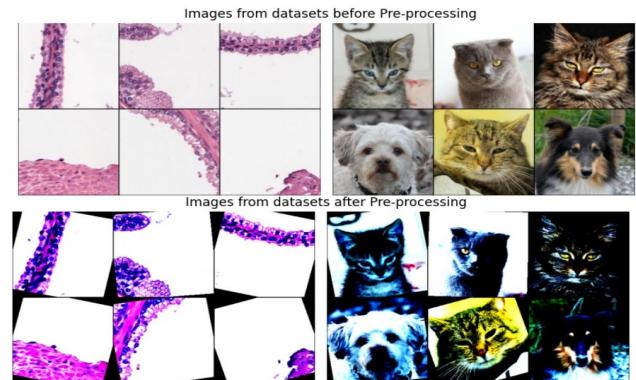


Figure 2. Task 2 datasets images before and after preprocessing

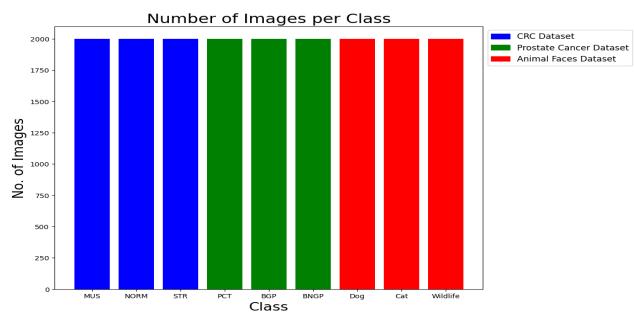


Figure 3. Statistics show classes balance of the datasets

cally ResNet-18, is chosen as the CNN architecture for its efficient design and improved gradient flow through residual connections [30]. ResNet-18 is an 18-layer network utilizing residual connections, consisting of an initial convolutional layer, followed by eight residual blocks, each comprising two 3x3 convolutional layers, and ending with a fully connected layer for classification. The first model, denoted as (CRC-Enc), is initialized by random weights. Con-

162  
163  
164  
165  
166  
167  
168  
169  
170  
171  
172  
173  
174  
175  
176  
177  
178  
179  
180  
181  
182  
183  
184  
185  
186  
187  
188  
189  
190  
191  
192  
193  
194  
195  
196  
197  
198  
199  
200  
201  
202  
203  
204  
205  
206  
207  
208  
209  
210  
211  
212  
213  
214  
215

versely, the second model, named (Imag-Enc), uses transfer learning as it is initialized with pre-trained weights from the ImageNet dataset. The output layer of both CRC-Enc and Imag-Enc is adapted to align with the number of labels present in the CRC dataset. Subsequently, training is conducted across all layers of these models. The training process employs the Adam optimizer, chosen to enhance the adaptability of the models to the underlying data distribution, thereby improving accuracy. The Adam optimizer is particularly well-suited for this purpose due to its ability to efficiently adjust learning rates during training, making it less sensitive to the initial choice of learning rate. This adaptability helps the model converge faster and achieve better solutions, particularly beneficial for complex and high-dimensional data. Various loss functions like Cross-Entropy Loss and Negative Log-Likelihood are explored. Cross-Entropy Loss is often preferred for classification tasks as it measures the dissimilarity between predicted probabilities and actual labels, encouraging the model to produce more confident and accurate predictions. A hyperparameter grid search is performed which optimizes the learning rate, batch size, and selection of the loss function. This optimization process involved systematic experimentation with various combinations of hyperparameter values in order to determine the most effective configuration for the machine learning model. Moreover, computational efficiency is an important factor in real-time applications. Accordingly, the computational workload is assessed by calculating the number of Floating-Point Operations Per Second (FLOPs) needed for inference. Our analytical findings reveal that the ResNet-18 models under examination consume approximately 1.82 billion FLOPs. Additionally, an average time per epoch of 50 seconds was calculated for training, which contributes to our understanding of the models' efficiency in practical settings. The experiments were conducted on Google Colab. Python 3.9 and PyTorch 2.0 were utilized with GPUs(Nvidia T4) to improve the training speed.

The second task encompasses feature extraction from two models, CRC-Enc and a ResNet18 model initialized with pre-trained weights from the ImageNet dataset, denoted as Pre-Img. This feature extraction is conducted on two distinct datasets which are prostate cancer and animal faces datasets. The trained models, CRC-Enc and Pre-Img, are used without their final classification layer. The objective is to investigate the transferability of feature representations in the context of transfer learning, with the presumption that a model's features are more adaptable to a new dataset when the domains of the original and target datasets exhibit close alignment. Since both colorectal and prostate cancer datasets relate to medical imaging, CRC-Enc exhibits superior performance in interpreting medical images. Conversely, Pre-Img maintains a broader feature

understanding, potentially enhancing its ability to recognize animal faces. Subsequently, the extracted features are employed for classification using ML techniques such as Support Vector Machine, K-Nearest Neighbors, and Random Forest. SVM finds the optimal hyperplane that maximizes the margin between different classes in a feature space. SVM is known for its effectiveness in handling high-dimensional feature spaces and finding optimal decision boundaries. KNN is a simple yet effective algorithm that assigns a class to a data point based on the majority vote from its 'k' nearest neighbors in the feature space. KNN effectively leverages the similarity between data points in well-represented features. It works well when data points are well-represented in features and exhibit local patterns. However, identifying the optimal 'k' value empowers the model to generalize effectively to previously unseen data. Random Forest is a powerful ensemble learning method that combines multiple decision trees to make predictions. It leverages the collective wisdom of these decision trees to provide robust and accurate classifications. In this project, these machine learning techniques are applied to assess their suitability for the classification tasks. SVM is chosen for its ability to handle high-dimensional data and complex decision boundaries, KNN is used for its simplicity and adaptability to localized patterns, and Random Forest is selected for its robustness and feature importance analysis capabilities. A comprehensive performance comparison among these three classifiers is conducted to evaluate their suitability for the specific classification tasks.

In both tasks, specific evaluation and assistance methods are employed to assess model performance and facilitate understanding. Firstly, a confusion matrix is utilized as a visual tool to represent the classification performance of the models. It provides insights by highlighting instances of misclassification, thus allowing for a deeper understanding of how the models are performing. Additionally, the confusion matrix enables the computation of crucial metrics such as precision, recall, and the F1 score, which are essential for quantifying the accuracy and effectiveness of the models in classifying data. Precision measures positive prediction accuracy, critical in contexts like medical diagnoses where false positives are costly and undesirable. Similarly, recall assesses the model's ability to find all actual positives. The F1 score offers a balanced performance assessment, especially useful in imbalanced class distributions or varying false positive and false negative costs. Furthermore, class representation and differentiation are visualized using t-distributed Stochastic Neighbor Embedding (t-SNE) applied to deep-layer activations [31, 32]. This technique reduces the dimensionality of features while preserving the relationships between data points. It shows how well the model distinguishes between different classes and how data clusters in the feature space. This gives insights into the

216  
217  
218  
219  
220  
221  
222  
223  
224  
225  
226  
227  
228  
229  
230  
231  
232  
233  
234  
235  
236  
237  
238  
239  
240  
241  
242  
243  
244  
245  
246  
247  
248  
249  
250  
251  
252  
253  
254  
255  
256  
257  
258  
259  
260  
261  
262  
263  
264  
265  
266  
267  
268  
269  
270  
271  
272  
273  
274  
275  
276  
277  
278  
279  
280  
281  
282  
283  
284  
285  
286  
287  
288  
289  
290  
291  
292  
293  
294  
295  
296  
297  
298  
299  
300  
301  
302  
303  
304  
305  
306  
307  
308  
309  
310  
311  
312  
313  
314  
315  
316  
317  
318  
319  
320  
321  
322  
323

324  
325  
326  
327  
328  
329  
330  
331  
332  
333  
334  
335  
336  
337  
338  
339  
340  
341  
342  
343  
344  
345  
346  
347  
348  
349  
350  
351  
352  
353  
354  
355  
356  
357  
358  
359  
360  
361  
362  
363  
364  
365  
366  
367  
368  
369  
370  
371  
372  
373  
374  
375  
376  
377

model's ability to separate classes effectively. Finally, the outcomes of this study offer insights, enabling engineers and scientists to make decisions regarding model selection and domain customization without the need for extensive training.

## D. Results

In Task 1, an extensive hyperparameter grid search was conducted to determine the optimal model configuration. The grid search encompassed a range of batch sizes, including 128, 64, and 32, as well as different learning rates spanning from 0.00001 to 0.05, with the training conducted over 10 epochs. The optimal model configuration was identified with a batch size of 32 and a learning rate of 0.0001. This configuration demonstrated smooth convergence within 10 epochs, as illustrated in the accuracy plot Fig. 4. Additionally, multiple loss functions, including Cross-Entropy Loss and Negative Log-Likelihood, were assessed. Cross-entropy loss resulted in a higher test accuracy (95.33%) compared to Negative Log-Likelihood (94.8%) with the chosen configuration setting, as evidenced by Fig. 11 in the supplementary appendix. Furthermore, the chosen combination resulted in achieving the minimum loss, as shown in Fig. 5 along with its confusion matrix in Fig. 6. Alternative configurations presented issues such as noise, slower convergence, or tendencies toward overfitting.

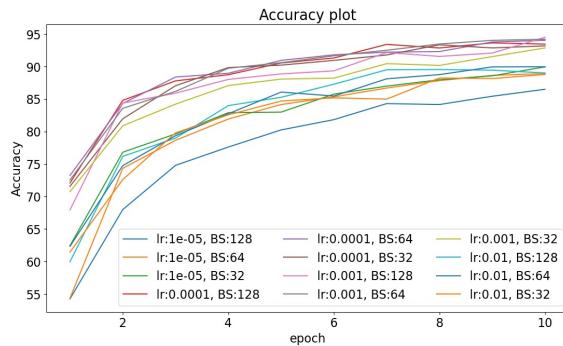


Figure 4. Accuracy plots for different configurations.

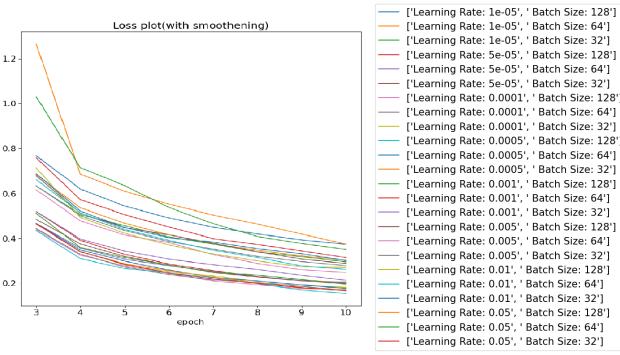


Figure 5. Loss function plots for different configurations.

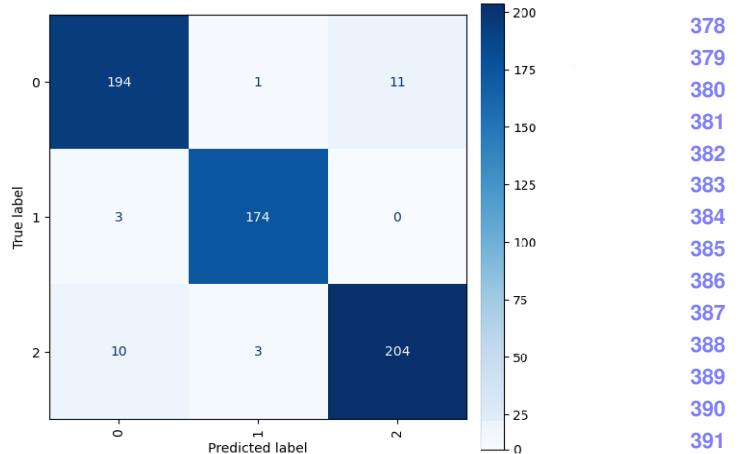


Figure 6. Confusion Matrix for BS=32, Lr=0.0001, Cross Entropy loss

The optimal configuration was employed for training both the CRC-Enc and Img-Enc models using the CRC dataset, and their performance is detailed in the initial section of Table 2. Notably, the Img-Enc pre-trained model outperforms the CRC-Enc model trained from scratch, benefiting from its prior knowledge of image features. Furthermore, fine-tuning Img-Enc with the CRC dataset significantly improves classification accuracy (99% for Img-Enc compared to 94% for CRC-Enc). A class separation visualization is displayed in Fig. 7, demonstrating effective class separation for both models, with Img-Enc exhibiting notably greater separation. Img-Enc's transfer learning approach, leveraging knowledge from a diverse dataset, contributed valuable low-level features, making it highly effective for cancer image classification. It is worth noting that CRC-Enc can potentially achieve performance matching that of Img-Enc when given enough training time. As a result, this emphasizes the notion that transfer learning accelerates convergence and enhances performance, highlighting the advantages of using pre-trained models for specialized tasks.

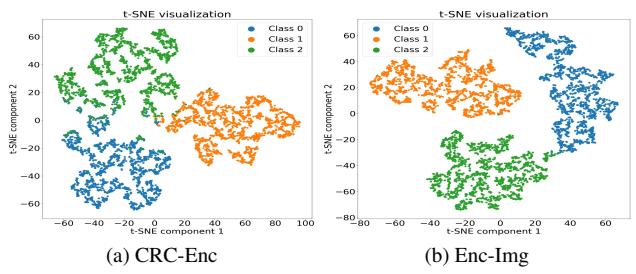


Figure 7. t-SNE plots of CRC dataset by the two models

Task 2 involved leveraging the feature extraction capabilities of both CRC-Enc and Pre-Img, which had been pretrained on ImageNet. Subsequently, a thorough assessment of their performance was conducted across

378  
379  
380  
381  
382  
383  
384  
385  
386  
387  
388  
389  
390  
391  
392  
393  
394  
395  
396  
397  
398  
399  
400  
401  
402  
403  
404  
405  
406  
407  
408  
409  
410  
411  
412  
413  
414  
415  
416  
417  
418  
419  
420  
421  
422  
423  
424  
425  
426  
427  
428  
429  
430  
431

432	Dataset	Model	F1-score	Precision	Recall	Test accuracy
433	NCT-CRC-HE-100K	CRC-Enc	0.95	0.95	0.95	94.83%
434		Imag-Enc	0.99	0.99	0.99	99.16%
435	After Feature Extraction	Classifier				
436	Prostate cancer	SVM	CRC-Enc	0.93	0.93	0.93
437		SVM	Pre-Img	0.96	0.96	95.58%
438		KNN	CRC-Enc	0.89	0.89	0.89
439		KNN	Pre-Img	0.93	0.94	0.93
440		RF	CRC-Enc	0.91	0.91	0.91
441		RF	Pre-Img	0.93	0.93	92.75%
442	Animal faces	SVM	CRC-Enc	0.68	0.68	0.68
443		SVM	Pre-Img	1.00	1.00	1.00
444		KNN	CRC-Enc	0.50	0.51	0.50
445		KNN	Pre-Img	0.99	0.99	0.99
446		RF	CRC-Enc	0.56	0.56	0.56
447		RF	Pre-Img	0.99	0.99	99.25%

Table 2. Detailed Results summary for Task 1 and 2

both prostate cancer and animal face datasets using SVM, KNN and RF as the chosen ML Classifiers. Both CRC-Enc and Pre-Img demonstrate high performance in classifying the prostate cancer dataset. Particularly, Pre-Img achieves slightly higher accuracy than CRC-Enc with SVM (95.58% for Pre-Img vs 93.08% for CRC-Enc) and KNN (93.08% vs. 88.75%) as in Table 2. These results are likely attributed to the diverse and extensive ImageNet dataset compared to the limited-size prostate cancer dataset. Additionally, t-SNE visualizations are provided in Fig. 8. The extracted features of both models before classification are represented in Fig.8-a and Fig.8-b. On the other hand, Fig.8 Fig.8-c and Fig.8-d revealed that both models achieved good class separation by SVM classifier and similarly for KNN classifier as in Fig.8-e and Fig.8-d. However, Pre-Img exhibited clearer feature separation for prostate cancer dataset, indicating its effectiveness in capturing relevant distinctions within the dataset.

In the context of animal face classification, Significant differences became apparent in the performance of feature representations produced by CRC-Enc and Pre-Img. The features adapted by CRC-Enc, initially designed for medical images, displayed diminished performance levels. In contrast, Pre-Img's more generalized feature representations achieved an impressive level of accuracy, particularly evident in the SVM classification, where it attained 99.58% accuracy compared to CRC-Enc's 67.5%. A similar trend was observed in KNN classification, with Pre-Img achieving 99.58% accuracy versus 67.5% for CRC-Enc. It is important to highlight that in all evaluated scenarios, SVM consistently demonstrated superior performance compared to KNN, even when the 'k' hyperparameter was meticulously optimized for KNN. This superiority can be attributed to SVM's capability to excel with limited data by establishing a global decision boundary that effectively generalizes from the available dataset. Conversely, KNN faces inherent challenges when dealing with high-dimensional feature spaces, often necessitating a larger dataset to attain reliable performance.

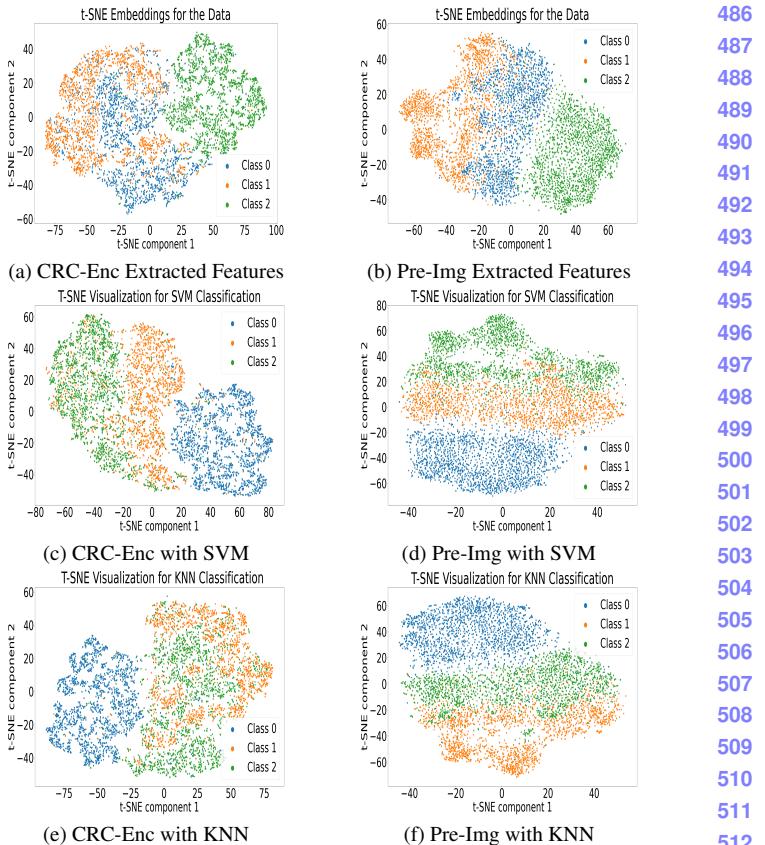
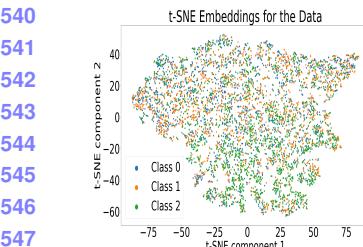


Figure 8. t-SNE visualizations present the CRC-Enc and Pre-Img extracted features, which are classified by SVM and KNN across the Prostate Cancer dataset.

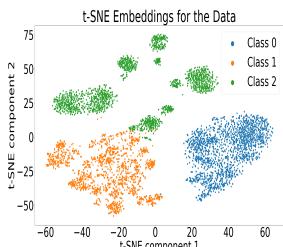
Additionally, insights gained from t-SNE visualizations, as shown in Fig. 9. The extracted features of both models before classification are represented in Fig.9-a and Fig.9-b. It is evident from Fig.9-d by SVM and Fig.9-f by KNN that Pre-Img demonstrated proficiency in effectively distinguishing between the classes within the Animal Faces dataset. Conversely, CRC-Enc features exhibited difficulty in adequately separating the dataset's distinct classes as shown in Fig.9-c by SVM classifier and Fig.9-e KNN classifier. These observations underscore the critical role played by feature extraction methods in influencing the model's performance in tasks across various domains. Pre-Img's capacity to generalize effectively, stemming from its training on a diverse range of images in the ImageNet dataset, facilitated superior performance in the animal face classification task, acknowledging the efficacy of leveraging pretrained models in diverse and specialized classification tasks.

Random Forest is known for its reliability and its capacity to assess important features. Given its power and robustness, it was intriguing to compare its performance to that of SVM and KNN. Pre-Img performed a bit better than CRC-Enc, achieving 93.50% accuracy compared to 91.58%

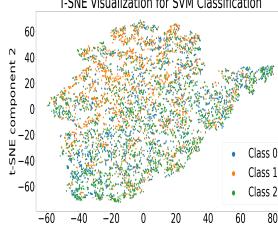
486  
487  
488  
489  
490  
491  
492  
493  
494  
495  
496  
497  
498  
499  
500  
501  
502  
503  
504  
505  
506  
507  
508  
509  
510  
511  
512  
513  
514  
515  
516  
517  
518  
519  
520  
521  
522  
523  
524  
525  
526  
527  
528  
529  
530  
531  
532  
533  
534  
535  
536  
537  
538  
539



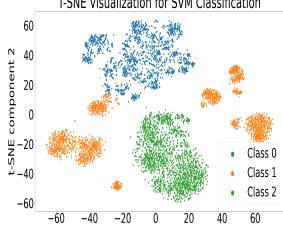
(a) CRC-Enc Extracted Features



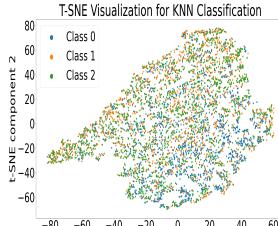
(b) Pre-Img Extracted Features



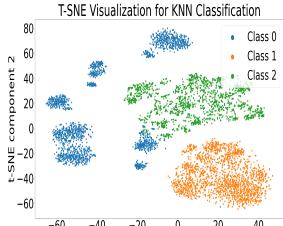
(c) CRC-Enc with SVM



(d) Pre-Img with SVM



(e) CRC-Enc with KNN



(f) Pre-Img with KNN

586

587

588

589

590

591

592

593

Figure 9. t-SNE visualizations present the CRC-Enc and Pre-Img extracted features, which are classified by an SVM and KNN across the Animal Faces dataset.

for CRC-Enc. Moreover, when dealing with animal faces, CRC-Enc's features didn't perform as well, achieving only 57.25% accuracy, while Pre-Img excelled with an accuracy of 99.16%. Random Forest excels in high-dimensional feature spaces, and its ability to create multiple decision trees and aggregate their results enables it to better capture complex relationships in the data, which is crucial in distinguishing animal faces. This explains why Random Forest outperformed KNN, which may encounter difficulties in high-dimensional spaces and when working with limited data.

In conclusion, the project's findings show that a CNN model with pre-trained weights performed significantly better than one trained from scratch on the CRC dataset, achieving an accuracy of 99% compared to 94%. This highlights the advantages of using pre-trained models, which lead to faster convergence and reduced training time. Additionally, pre-trained feature extractors from large, diverse datasets were more effective than those trained on medical datasets, resulting in better feature extraction, although their performance may decrease on datasets from diverse

domains. The results provide strong evidence that the success of transfer learning depends on the similarity between the pre-trained and target datasets. Transfer learning tends to be more successful when the two datasets closely resemble each other, but its effectiveness tends to diminish when there is significant dissimilarity between them. Furthermore, the choice of machine learning model greatly impacts performance, with SVM consistently outperforming KNN and RF, achieving significantly higher accuracy rates, with an average improvement of approximately 30% in domain-specific tasks. Overall, this project demonstrates the effectiveness of transfer learning, especially when utilizing pre-trained models for specialized tasks.

594  
595  
596  
597  
598  
599  
600  
601  
602  
603  
604  
605  
606  
607  
608  
609  
610  
611  
612  
613  
614  
615  
616  
617  
618  
619  
620  
621  
622  
623  
624  
625  
626  
627  
628  
629  
630  
631  
632  
633  
634  
635  
636  
637  
638  
639  
640  
641  
642  
643  
644  
645  
646  
647

648  
649  
650  
651  
652  
653  
654  
655  
656  
657  
658  
659  
660  
661  
662  
663  
664  
665  
666  
667  
668  
669  
670  
671  
672  
673  
674  
675  
676  
677  
678  
679  
680  
681  
682  
683  
684  
685  
686  
687  
688  
689  
690  
691  
692  
693  
694  
695  
696  
697  
698  
699  
700  
701

## References

- [1] Yanan Sun, Bing Xue, Mengjie Zhang, and Gary G. Yen. Evolving deep convolutional neural networks for image classification. *IEEE Transactions on Evolutionary Computation*, 24(2):394–407, 2020. 1
- [2] Can Taylan Sari and Cigdem Gunduz-Demir. Unsupervised Feature Extraction via Deep Learning for Histopathological Classification of Colon Tissue Images. *IEEE transactions on medical imaging*, 38(5):1139–1149, May 2019. 1
- [3] Sinno Jialin Pan and Qiang Yang. A survey on transfer learning. *IEEE Transactions on Knowledge and Data Engineering*, 22(10):1345–1359, 2010. 1
- [4] Bratislav Predić, Uroš Vukić, Muzafer Saračević, Darjan Karabašević, and Dragiša Stanujkić. The Possibility of Combining and Implementing Deep Neural Network Compression Methods. *Axioms*, 11(5):229, May 2022. 1
- [5] Wenyi Teng, Ni Wang, Huihui Shi, Yuchan Liu, and Jing Wang. Classifier-Constrained Deep Adversarial Domain Adaptation for Cross-Domain Semisupervised Classification in Remote Sensing Images. *IEEE Geoscience and Remote Sensing Letters*, 17(5):789–793, May 2020. Conference Name: IEEE Geoscience and Remote Sensing Letters. 1
- [6] Fabian Schenkel and Wolfgang Middelmann. Domain Adaptation for Semantic Segmentation Using Convolutional Neural Networks. In *IGARSS 2019 - 2019 IEEE International Geoscience and Remote Sensing Symposium*, pages 728–731, July 2019. ISSN: 2153-7003. 1
- [7] Timothy Hospedales, Antreas Antoniou, Paul Micaelli, and Amos Storkey. Meta-Learning in Neural Networks: A Survey. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 44(09):5149–5169, September 2022. Publisher: IEEE Computer Society. 1
- [8] Hojjat Salehinejad and Shahrokh Valaei. EDropout: Energy-Based Dropout and Pruning of Deep Neural Networks. *IEEE Transactions on Neural Networks and Learning Systems*, 33(10):5279–5292, October 2022. Conference Name: IEEE Transactions on Neural Networks and Learning Systems. 1
- [9] Zhang Chiliang, Hu Tao, Guan Yingda, and Ye Zuochang. Accelerating Convolutional Neural Networks with Dynamic Channel Pruning. In *2019 Data Compression Conference (DCC)*, pages 563–563, March 2019. ISSN: 2375-0359. 1
- [10] Thaker Pragnesh and Biju R. Mohan. Compression of Convolution Neural Network Using Structured Pruning. In *2022 IEEE 7th International conference for Convergence in Technology (I2CT)*, pages 1–5, April 2022. 1
- [11] Hsing-Hung Chou, Ching-Te Chiu, and Yi-Ping Liao. Deep Neural Network Compression with Knowledge Distillation Using Cross-Layer Matrix, KL Divergence and Offline Ensemble. In *2020 Asia-Pacific Signal and Information Processing Association Annual Summit and Conference (APSIPA ASC)*, pages 71–75, December 2020. ISSN: 2640-0103. 1
- [12] Bei Liu, Haoyu Wang, Zhengyang Chen, Shuai Wang, and Yanmin Qian. Self-Knowledge Distillation via Feature Enhancement for Speaker Verification. In *ICASSP 2022 - 2022 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pages 7542–7546, May 2022. ISSN: 2379-190X. 1
- [13] Duc-Quang Vu, Ngan Le, and Jia-Ching Wang. Teaching Yourself: A Self-Knowledge Distillation Approach to Action Recognition. *IEEE Access*, 9:105711–105723, 2021. Conference Name: IEEE Access. 1
- [14] Akhilesh Kumar Sharma, Shamik Tiwari, Gaurav Aggarwal, Nitika Goenka, Anil Kumar, Prasun Chakrabarti, Tulika Chakrabarti, Radomir Gono, Zbigniew Leonowicz, and Michał Jasiński. Dermatologist-level classification of skin cancer using cascaded ensembling of convolutional neural network and handcrafted features based deep neural network. *IEEE Access*, 10:17920–17932, 2022. 1
- [15] Arpit Jain, Anurag Gupta, Alok Singh Sengar, Ratnesh Kumar Shukla, and Abhishek Jain. Application of Deep Learning for Image Sequence Classification. In *2021 10th International Conference on System Modeling & Advancement in Research Trends (SMART)*, pages 280–284, December 2021. ISSN: 2767-7362. 1
- [16] Alex Krizhevsky, Ilya Sutskever, and Geoffrey E. Hinton. ImageNet classification with deep convolutional neural networks. *Communications of the ACM*, 60(6):84–90, May 2017. 1
- [17] Karen Simonyan and Andrew Zisserman. Very Deep Convolutional Networks for Large-Scale Image Recognition, April 2015. arXiv:1409.1556 [cs]. 1
- [18] Ashish Tripathi, Anuradha Misra, Kuldeep Kumar, and Brijesh Kumar Chaurasia. Optimized machine learning for classifying colorectal tissues. *SN Computer Science*, 4, 9 2023. 1
- [19] Qasem Abu Al-Haija and Adeola Adebanjo. Breast Cancer Diagnosis in Histopathological Images Using ResNet-50 Convolutional Neural Network. In *2020 IEEE International IOT, Electronics and Mechatronics Conference (IEMTRON-ICS)*, pages 1–7, September 2020. 1
- [20] Abeer Saber, Mohamed Sakr, Osama M. Abo-Seida, Arabi Keshk, and Huijing Chen. A novel deep-learning model for automatic detection and classification of breast cancer using the transfer-learning technique. *IEEE Access*, 9:71194–71209, 2021. 1
- [21] Raphael M. Kronberg, Lena Haeberle, Melanie Pfaus, Haifeng C. Xu, Karina S. Krings, Martin Schlenzog, Tilman Rau, Aleksandra A. Pandry, Karl S. Lang, Irene Esposito, and Philipp A. Lang. Communicator-Driven Data Preprocessing Improves Deep Transfer Learning of Histopathological Prediction of Pancreatic Ductal Adenocarcinoma. *Cancers*, 14(8):1964, January 2022. Number: 8 Publisher: Multidisciplinary Digital Publishing Institute. 1
- [22] Bratislav Predić, Uroš Vukić, Muzafer Saračević, Darjan Karabašević, and Dragiša Stanujkić. The Possibility of Combining and Implementing Deep Neural Network Compression Methods. *Axioms*, 11(5):229, May 2022. Number: 5 Publisher: Multidisciplinary Digital Publishing Institute. 1
- 702  
703  
704  
705  
706  
707  
708  
709  
710  
711  
712  
713  
714  
715  
716  
717  
718  
719  
720  
721  
722  
723  
724  
725  
726  
727  
728  
729  
730  
731  
732  
733  
734  
735  
736  
737  
738  
739  
740  
741  
742  
743  
744  
745  
746  
747  
748  
749  
750  
751  
752  
753  
754  
755

## COMP 6321 Submission #E. Concordia University - Fall 2023

- 756 [23] Raul Victor Medeiros da Nóbrega, Solon Alves Peixoto,  
757 Suane Pires P. da Silva, and Pedro Pedrosa Rebouças Filho.  
758 Lung Nodule Classification via Deep Transfer Learning in  
759 CT Lung Images. In *2018 IEEE 31st International Symposium on Computer-Based Medical Systems (CBMS)*, pages  
760 244–249, June 2018. ISSN: 2372-9198. 1
- 761 [24] Kaiming He, Ross Girshick, and Piotr Dollar. Rethinking  
762 ImageNet Pre-Training. In *2019 IEEE/CVF International Conference on Computer Vision (ICCV)*, pages 4917–4926,  
763 October 2019. ISSN: 2380-7504. 1
- 764 [25] Indranil Ray, Geetank Raipuria, and Nitin Singhal. Rethinking  
765 ImageNet Pre-training for Computational Histopathology.  
766 In *2022 44th Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC)*,  
767 pages 3059–3062, July 2022. ISSN: 2694-0604. 1, 2
- 768 [26] Simon Kornblith, Jonathon Shlens, and Quoc V. Le. Do Better  
769 ImageNet Models Transfer Better? In *2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*,  
770 pages 2656–2666, June 2019. ISSN: 2575-7075. 1
- 771 [27] Jakob Nikolas Kather, Niels Halama, and Alexander Marx.  
772 100,000 histological images of human colorectal cancer and  
773 healthy tissue. Zenodo, April 2018. 2
- 774 [28] Yuri Tolkach. Datasets digital pathology and artifacts, part  
775 1. Zenodo, May 2021. 2
- 776 [29] Yunjey Choi, Youngjung Uh, Jaejun Yoo, and Jung-Woo Ha.  
777 Stargan v2: Diverse image synthesis for multiple domains.  
778 In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2020. 2
- 779 [30] Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun.  
780 Deep Residual Learning for Image Recognition. In *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*,  
781 pages 770–778, Las Vegas, NV, USA, June 2016. IEEE. 2
- 782 [31] Jakob Nikolas Kather, Johannes Krisam, Pornpimol  
783 Charoentong, Tom Luedde, Esther Herpel, Cleo Aron Weis,  
784 Timo Gaiser, and Alexander Marx. Predicting survival from  
785 colorectal cancer histology slides using deep learning: A retrospective multicenter study. *PLoS MTripathi2023edicine*,  
786 16, 2019. 3
- 787 [32] David Ahmedt-Aristizabal, Mohammad Ali Armin, Simon  
788 Denman, Clinton Fookes, and Lars Petersson. A survey on  
789 graph-based deep learning for computational histopathology.  
790 *Computerized Medical Imaging and Graphics*, 95, 1 2022. 3
- 791 [33] Michael J. Zitnick, Ali Farhadi, and Deva Ramanan. Deep  
792 multi-view learning for multi-person 3D reconstruction. In  
793 *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*,  
794 pages 1021–1028, June 2014. IEEE. 3
- 795 [34] Michael J. Zitnick, Ali Farhadi, and Deva Ramanan. Deep  
796 multi-view learning for multi-person 3D reconstruction. In  
797 *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*,  
798 pages 1021–1028, June 2014. IEEE. 3
- 799 [35] Michael J. Zitnick, Ali Farhadi, and Deva Ramanan. Deep  
800 multi-view learning for multi-person 3D reconstruction. In  
801 *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*,  
802 pages 1021–1028, June 2014. IEEE. 3
- 803 [36] Michael J. Zitnick, Ali Farhadi, and Deva Ramanan. Deep  
804 multi-view learning for multi-person 3D reconstruction. In  
805 *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*,  
806 pages 1021–1028, June 2014. IEEE. 3
- 807 [37] Michael J. Zitnick, Ali Farhadi, and Deva Ramanan. Deep  
808 multi-view learning for multi-person 3D reconstruction. In  
809 *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*,  
810 pages 1021–1028, June 2014. IEEE. 3

864

## E. Supplementary Material

865

866

867

868

869

870

871

872

873

874

875

876

877

878

879

880

881

882

883

884

885

886

887

888

889

890

891

892

893

894

895

896

897

898

899

900

901

902

903

904

905

906

907

908

909

910

911

912

913

914

915

916

917

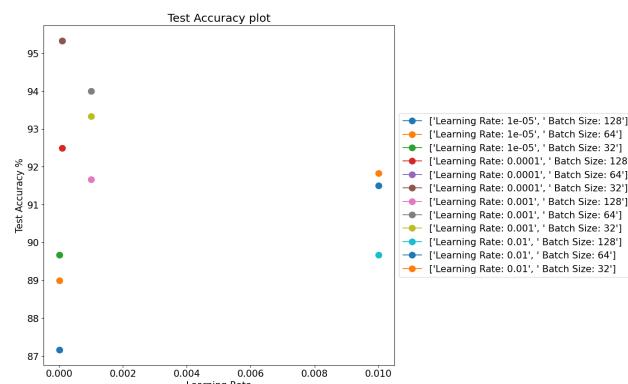


Figure 10. Accuracy plot for different configurations.

```

Learning Rate: 0.0001, Batch Size: 32 , Loss function: Cross Entropy Loss
Optimizer: Adam Optimizer Device: cuda
#####
Training Finished in 512.7450790405273 seconds #####
#####
Training timer per epoch is 51.27450790405273 seconds #####
Test Accuracy of the model on the 600 test images: 95.33333333333334 %
precision recall f1-score support
0 0.94 0.94 0.94 286
1 0.98 0.98 0.98 177
2 0.95 0.94 0.94 217
accuracy 0.95 0.95 0.95 600
macro avg 0.95 0.95 0.95 600
weighted avg 0.95 0.95 0.95 600

Learning Rate: 0.0001, Batch Size: 32 , Loss function: Negative Log-Likelihood
Optimizer: Adam Optimizer Device: cuda
#####
Training Finished in 513.411628484726 seconds #####
#####
Training timer per epoch is 51.34116284847259 seconds #####
Test Accuracy of the model on the 600 test images: 94.83333333333334 %
precision recall f1-score support
0 0.90 0.96 0.93 190
1 0.97 0.99 0.98 192
2 0.97 0.90 0.93 218
accuracy 0.95 0.95 0.95 600
macro avg 0.95 0.95 0.95 600
weighted avg 0.95 0.95 0.95 600

```

Figure 11. Comparison of loss function for best combination in task 1.

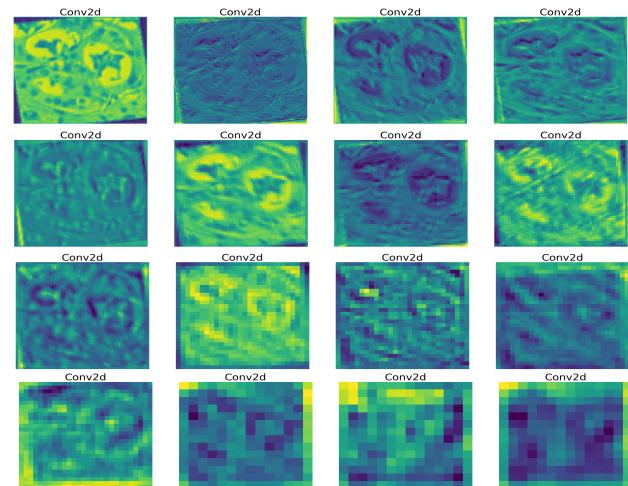
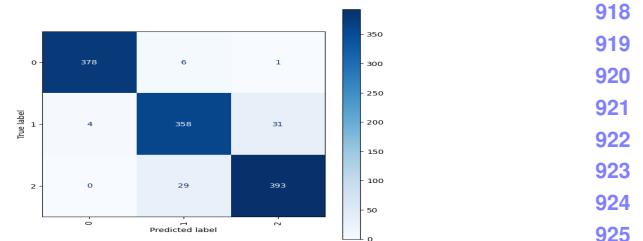
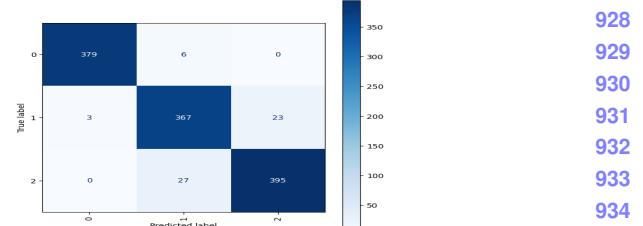


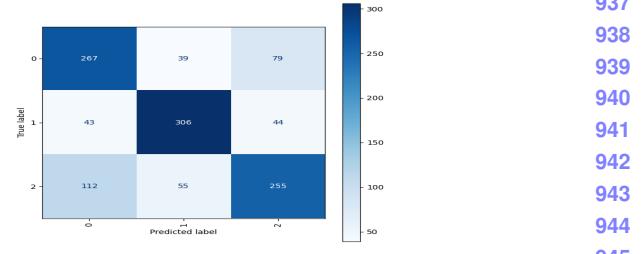
Figure 12. Features Map of Conv2 layer of CRC-Enc feature extractor model



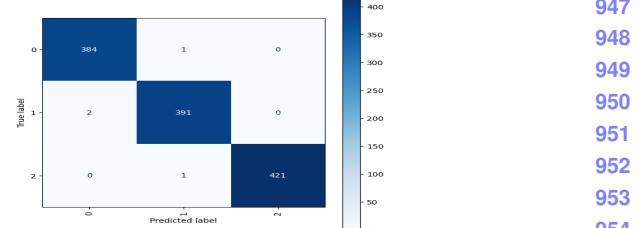
(a) Prostate Cancer features extracted using CRC-Enc Classified by SVM



(b) Prostate Cancer features extracted using Pre-Img Classified by SVM



(c) Animal Faces features extracted using CRC-Enc Classified by SVM



(d) Animal Faces features extracted using Pre-Img Classified by SVM

Figure 13. Confusion matrix of extracted features from Prostate Cancer and Animal Faces datasets Classified by SVM

972

973

974

975

976

977

978

979

980

981

982

983

984

985

986

987

988

989

990

991

992

993

994

995

996

997

998

999

1000

1001

1002

1003

1004

1005

1006

1007

1008

1009

1010

1011

1012

1013

1014

1015

1016

1017

1018

1019

1020

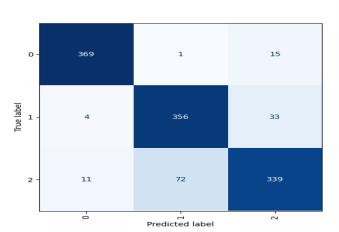
1021

1022

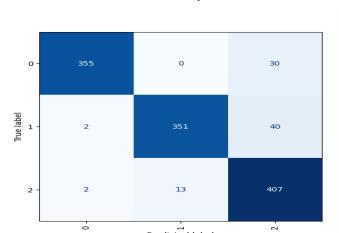
1023

1024

1025



(a) Prostate Cancer features extracted using CRC-Enc Classified by KNN



(b) Prostate Cancer features extracted using Pre-Img Classified by KNN



(c) Animal Faces features extracted using CRC-Enc Classified by KNN



(d) Animal Faces features extracted using Pre-Img Classified by KNN

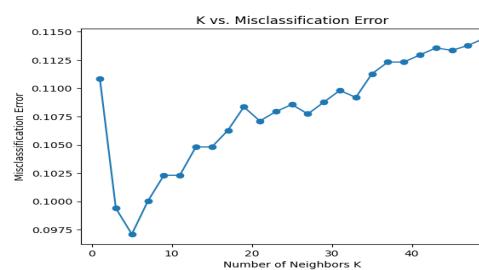
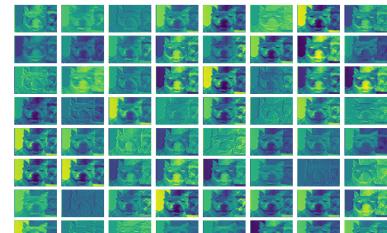


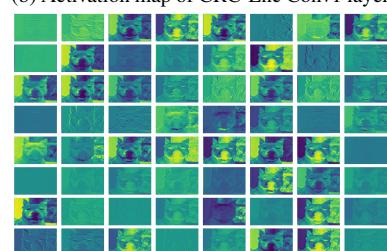
Figure 15. Optimizing Hyperparameter 'K' for KNN training.



(a) Input Image of Animal Faces

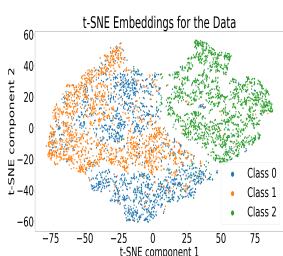


(b) Activation map of CRC-Enc Conv1 layer

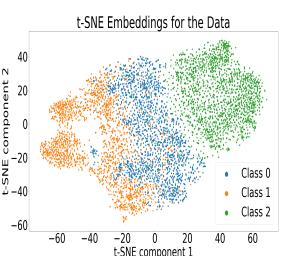


(c) Activation map of Pre-Img Conv1 layer

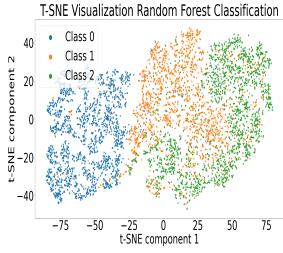
Figure 16. visualizations the activations of CRC-Enc and Pre-Img Conv1 Layer of Animal Faces dataset



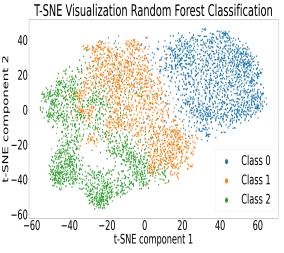
(a) CRC-Enc Extracted Features



(b) Pre-Img Extracted Features



(c) CRC-Enc with RF



(d) Pre-Img with RF

Figure 17. t-SNE visualizations present the CRC-Enc and Pre-Img extracted features, which are classified by Random Forest across the Prostate Cancer dataset.

1026

1027

1028

1029

1030

1031

1032

1033

1034

1035

1036

1037

1038

1039

1040

1041

1042

1043

1044

1045

1046

1047

1048

1049

1050

1051

1052

1053

1054

1055

1056

1057

1058

1059

1060

1061

1062

1063

1064

1065

1066

1067

1068

1069

1070

1071

1072

1073

1074

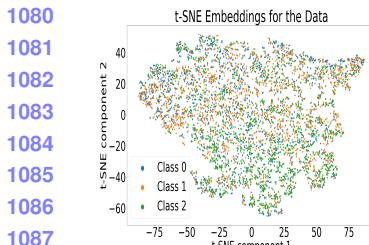
1075

1076

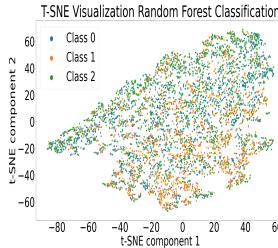
1077

1078

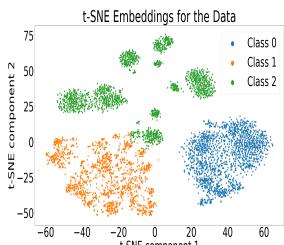
1079



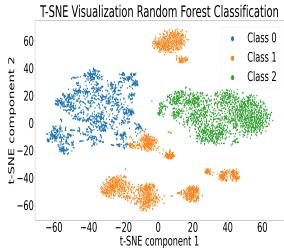
(a) CRC-Enc Extracted Features



(c) CRC-Enc with RF

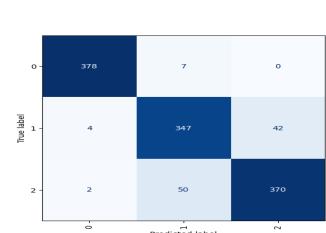


(b) Pre-Img Extracted Features

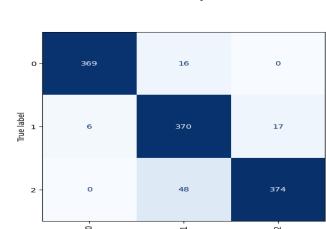


(d) Pre-Img with RF

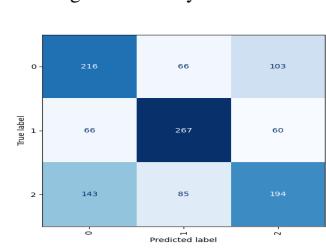
Figure 18. t-SNE visualizations present the CRC-Enc and Pre-Img extracted features, which are classified by Random Forest across the Animal Faces dataset.



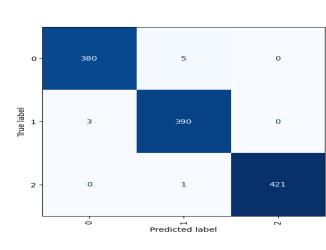
(a) Prostate Cancer features extracted using CRC-Enc Classified by RF



(b) Prostate Cancer features extracted using Pre-Img Classified by RF



(c) Animal Faces features extracted using CRC-Enc Classified by RF



(d) Animal Faces features extracted using Pre-Img Classified by RF

Figure 19. Confusion matrix of extracted features from Prostate Cancer and Animal Faces datasets Classified by Random Forest

1134

1135

1136

1137

1138

1139

1140

1141

1142

1143

1144

1145

1146

1147

1148

1149

1150

1151

1152

1153

1154

1155

1156

1157

1158

1159

1160

1161

1162

1163

1164

1165

1166

1167

1168

1169

1170

1171

1172

1173

1174

1175

1176

1177

1178

1179

1180

1181

1182

1183

1184

1185

1186

1187