JOURNAL OF
EMERGING INVESTIGATORS

Article

Transfer learning and data augmentation in osteosarcoma cancer detection

Jason Chu[1], Sarah Khan[2]
[1] Monta Vista High School, Cupertino, California
[2] AIClub, Santa Clara, California

SUMMARY

Osteosarcoma is a type of bone cancer that affects young adults and children. Early diagnosis of osteosarcoma is crucial to successful treatment. The current methods of diagnosis, which include imaging tests and biopsy, are time consuming and prone to human error. Hence, we used deep learning to extract patterns and detect osteosarcoma from histological images. We hypothesized that the combination of two different technologies (transfer learning and data augmentation) would improve the efficacy of osteosarcoma detection in histological images. The dataset used for the study consisted of histological images for osteosarcoma and was quite imbalanced as it contained very few images with tumors. Since transfer learning uses existing knowledge for the purpose of classification and detection, we hypothesized it would be proficient on such an imbalanced dataset. To further improve our learning, we used data augmentation to include variations in the dataset. We further evaluated the efficacy of different convolutional neural network models on this task. We obtained an accuracy of 91.18% using the transfer learning model MobileNetV2 as the base model with various geometric transformations, outperforming the state-of-the-art convolutional neural network based approach.

INTRODUCTION

Cancer is characterized by abnormal cell growth when cell checkpoints and growth inhibitors fail to function. Osteosarcoma is a type of malignant bone cancer, most often plaguing young adults and children between the ages of 10 and 20 (1). It is the most prevalent malignant bone tumor in children (2). Between 2015 and 2019, cancer in the bones and joints was the second deadliest (behind leukemia) amongst adolescents aged 15-19 (3). However, osteosarcoma is not limited to adolescents; 1 in 10 occurrences are detected in people older than 60 (4). About 75% of patients can survive if the condition has not spread to other parts of the body; therefore, early diagnosis is critical. Symptoms include dull aching or pain in the bone or joint, which can also form swelling due to the growth of the tumor, progressing to movement deficiencies and spreading to lungs or adjacent bones (1).

Due to the seemingly harmless nature of the early symptoms, early diagnosis can be affected, which makes the cancer more difficult to treat. Often people are not diagnosed with osteosarcoma in the early symptoms and stages because they may not realize their symptoms indicate a serious issue. Diagnosis of osteosarcoma may require a family physician, orthopedic oncologist, medical oncologist, radiologist, and pathologist (5). Current methods of diagnosis include imaging tests and biopsy. Imaging tests require examination by doctors and are subject to error since the diagnosis from images can be very subjective and variances between patients can make correct identification more difficult (6). A biopsy involves sampling cells and testing them in a laboratory for abnormalities. This is required to confirm a diagnosis. The process of detecting osteosarcoma is time-consuming; however, deep learning models make the process more efficient. Deep learning models only need to be trained on large data sets once, and then they can be used directly on patients' data, saving diagnosis time. In contrast to deep learning models, human diagnosis involves more room for error and takes more time (6).

We used one of the popular methods, transfer learning (7), which is a method that focuses on storing knowledge from one problem and applying it to another (**Figure 1**). It is a popular method in the deep learning space and can train deep neural networks with a small amount of data. Transfer learning is often used to solve the problem of insufficient training data by transferring knowledge from the source domain, which contains ample data points, to the target domain, which has fewer data points. This helped us detect osteosarcoma through histological images. Transfer learning enabled us to share the knowledge obtained from the best current models for classifying images.

Data augmentation tries to create artificial data training points to enhance the size and quality of data by adding slightly

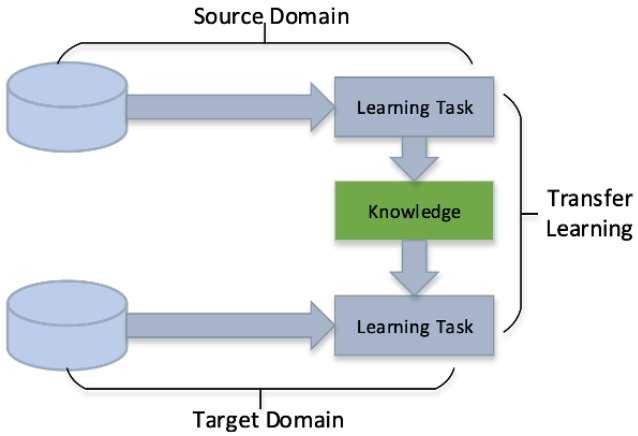

Figure 1: The process of transfer learning. The information learned from the task of the source domain (for example, ImageNet was the source domain in this study) is applied to learning the target domain. This is the process of transfer learning.

different data points. It is useful specifically when there is a scarcity of training data points and is very easily applicable to classification tasks such as object recognition, speech recognition, and image classification. Over the years, many data augmentation methods have been devised for numerous input types such as images, speech, text, and signals (14-17). Data augmentation is also a regularization technique that can be used to reduce overfitting (18) and improve generalization of deep learning models (18).

Data augmentation can be classified into two techniques: transforming input data points to create new points and learning dataset distribution to generate new synthetic data points. As there was a lack of images in the dataset used in this study, we decided to use data augmentation to increase the number of training images. We focused on image transformation techniques such as rotation, zoom, and flip.

We hypothesized that transfer learning techniques from deep neural networks could be effectively used to detect osteosarcoma from imbalanced datasets using data augmentation and would outperform other approaches. We used data augmentation and transfer learning to train a model from the dataset which consisted of 1144 images from 50 patients labeled Non-Tumor, Viable Tumor, and Non-Viable Tumor and obtained 91.18% accuracy using our process (19) (**Figure 2**). Our technique also outperformed a convolutional neural network (CNN) based approach that did not use transfer learning (9). Convolutional Neural Networks (CNNs) are a type of deep learning neural network commonly used in image and video recognition and classification tasks. It uses a process called convolution to extract features from the input images, and then applies functions to produce a classification output.

In this paper, we hypothesized that the combination of two different technologies, transfer learning and data augmentation, would improve the efficacy of osteosarcoma detection in histological images (7)(8). We evaluated this hypothesis via experimentation with an existing dataset, application of these techniques, and training evaluation using several state-of-the-art convolutional neural networks. In each case we evaluated our results against a base study done previously which utilized the same dataset and reported validation (testing accuracies) between 75-86% using various forms of oversampling and custom-designed CNNs (9). Using existing CNNs and additional techniques of transfer learning and data augmentation, we were able to outperform these prior results.

RESULTS

Following the hypothesis of using techniques of transfer learning and data augmentation to enhance the diagnosis of osteosarcoma, our first step was to evaluate each model using only transfer learning without augmentation. With transfer learning, we tried to exploit what has been learned in one task to improve generalization in another by transferring the weights of the model that has learned task “a” to task “b”. Hence, instead of starting the learning process from scratch, we started with the patterns learned from solving another task. We used pre-trained early and middle layers and only trained the latter layers. For this task, we evaluated four specific transfer learning convolutional neural networks – MobileNetV2, ResNet50V2, Xception, and InceptionResNetV2 – which are all pre-trained deep learning models used for image classification tasks (20-22). In each case the models used transfer learning, starting with a preloaded model trained on ImageNet, an open source large dataset widely used in computer vision research of about 20000 categories and over 14 million images. Each of these models has its own unique architecture and strengths. MobileNetV2 is a lightweight model that was designed for efficient use on mobile devices and edge computing. ResNet50V2 is a more powerful model that uses residual connections to train deeper networks. Xception is an architecture that combines depth-wise separable convolutions and is designed to be more accurate than MobileNetV2 but less computationally expensive than InceptionV3. InceptionResNetV2 is a hybrid model that combines the strengths of Inception and ResNet. It combines these two architectures to help improve the model’s accuracy while still being computationally efficient.

We used accuracy as our prime method of evaluation for each model. Accuracy is defined as the percentage of correct predictions for the test/validation data. It can be calculated by dividing the number of correct predictions by the number of total predictions. We obtained a range of average accuracies by testing different models at varying learning rates and epochs (**Figure 3**). Learning rate is the rate at which the model tries to minimize its loss. An epoch is a pass across

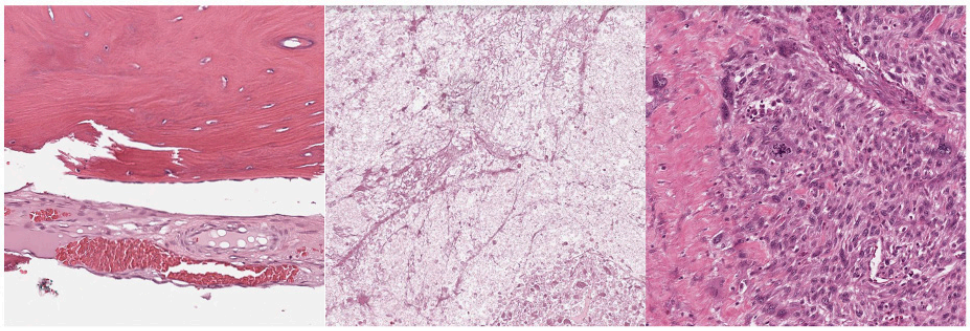

Figure 2: Examples of images of each class of osteosarcoma tissue from the dataset. The stained histological images shown are from the classes: Non-Tumor, Non-Viable Tumor, and Viable Tumor. Non-Tumor refers to a tissue that does not contain tumor, Non-Viable Tumor refers to tumor that has undergone necrosis, and Viable Tumor refers to a tumor capable of spreading.

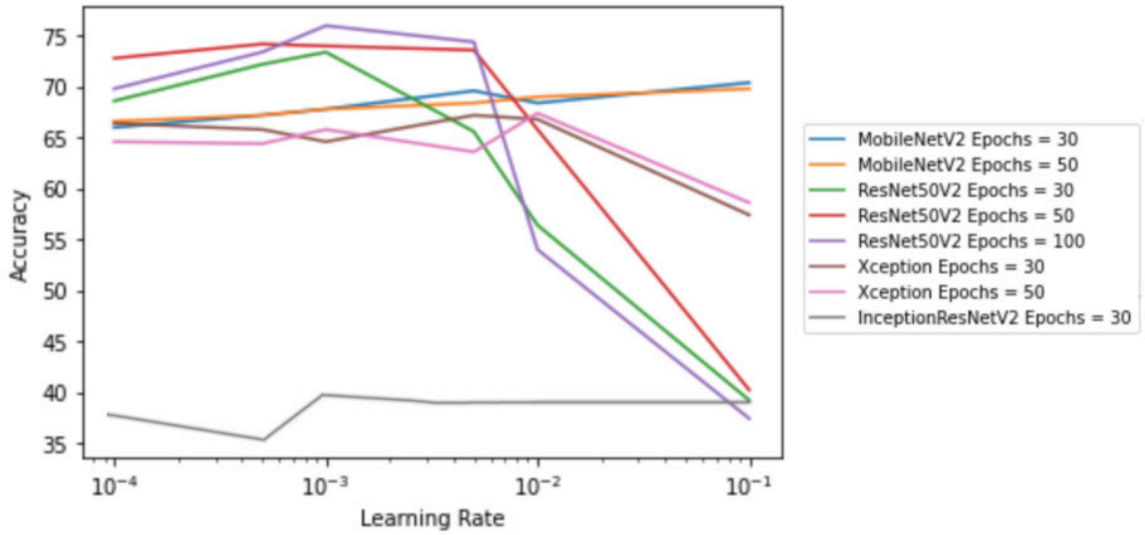

Figure 3: Average accuracies reached by different convolutional neural network models at varying learning rates and epochs. The transfer learning models MobileNetV2, ResNet50V2, Xception, and InceptionResNetV2 were tested at different epochs and learning rates. The accuracy (out of 100) at each learning rate is a calculated average accuracy of 5 trials.

the complete training dataset while training the model. ResNet50V2 consistently outperformed all the other network architectures. The second-best performing model was MobileNetV2. All models performed better at lower learning rates, with the exception of MobileNetV2, which appeared to improve in accuracy at higher learning rates. In summary, without data augmentation, the highest accuracy obtained was 79%, obtained by ResNet50V2.

While we obtained good results, we saw that the results did not uniformly improve over our reference results in paper (9) (75-86%). The next step was to explore data augmentation. We focused on ResNetV2 and MobileNetV2 because their performance was more consistent across different learning rates and epochs. Another reason for selecting these two models is the large difference in the number of trainable parameters. We used various geometric transformation techniques such as rotation, zoom, and flip for augmenting the data. We first created a copy of the dataset so that the original images and transformed images could both be kept. Then we used a combination of both for training the models. For the image transformation of random rotate, random rotation by an argument "x" rotates the image by any random number of degrees between -x° to x°. The results were tested up to 30 degrees, which generally reached the highest accuracy. Horizontal flip flips the image horizontally when set to true. Finally, the zoom image transformation by a factor of "x" will zoom the image by any value between 1-x and 1+x. Zooming by values of 0.5, 1, and 1.5 were tested, with 1.5 reaching the highest accuracy. This meant that the images in the final accuracy model were magnified by any factor between 0.5 and 2.5, a relatively large range. Results of some of the important experiments are listed (**Table 1**).

In this set of experiments, we observed that it is possible for MobileNetV2 to outperform ResNet50V2 when the data is suitably augmented. In fact, as more types of augmentation were added, the accuracy of MobileNetV2 inched further upward. We also observed that the epochs did not materially change the results. In sum, the best accuracy obtained was 91.18% by MobileNetV2 with data augmented by rotation, flip and zoom. The improvement by adding data augmentation was over 12% (best case of 79% without data augmentation compared to the best case of 91% with data augmentation). This result outperformed our reference base of 75-86%.

We then sought to minimize overfitting of our model by examining the training and validation accuracy as the number of epochs increases (**Figure 4**). Overfitting (where the neural network begins to memorize the training data) can be indicated by a training accuracy that continues to increase as the validation accuracy drops. While larger datasets (as provided by data augmentation) can reduce the likelihood of overfitting, our results demonstrated that even then the models could overfit if given enough training epochs. To avoid this issue, we reduced the number of epochs and reported results for the best performing validation accuracy as achieved prior to overfitting.

Augmentation	Model	Epochs	Learning Rate	Accuracy
Random rotate 30	ResNet50V2	30	0.001	80.39%
Random rotate 30	MobileNetV2	30	0.001	88.24%
Random rotate 30, Horizontal Flip	ResNet50V2	30	0.001	78.43%
Random rotate 30, Horizontal Flip	MobileNetV2	30	0.0001	87.25%
Random rotate 30, Horizontal Flip, Zoom 0.5	MobileNetV2	10	0.0001	89.22%
Random rotate 30, Horizontal Flip, Zoom 1	MobileNetV2	50	0.0001	89.22%
Random rotate 30, Horizontal Flip, Zoom 1.5	MobileNetV2	30	0.0001	91.18%

Table 1: The accuracies reached with varying epochs, learning rates, and geometric transformations for MobileNetV2 and ResNet50V2. Random rotation by "x" will rotate by any number of degrees between -x° to x°. The results were tested up to 30 degrees. Horizontal flip flips the image horizontally when set to true. Finally, the zoom image transformation by a factor of "x" will zoom the image by any value between 1-x and 1+x. Zooming by values of 0.5, 1, and 1.5 were tested, with 1.5 reaching the highest accuracy.

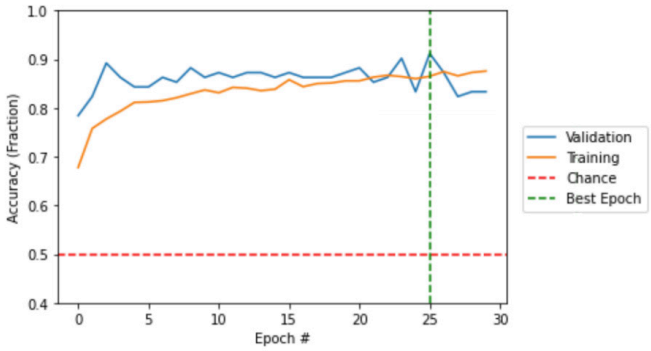

Table 4: The validation and training accuracy at each epoch of training the highest accuracy MobileNetV2 model. The validation and training accuracy reached in percentage at each epoch for the highest performing run of MobileNetV2. The best epoch reached an accuracy of 91.18% (vertical green line). Overfitting can be observed at 25 epochs.

As a final result, for the best performing model (MobileNetV2), we visualized the confusion matrix (**Figure 5**). A confusion matrix is a performance measurement tool for machine learning classification problems where the output can be two or more classes. It can be used for calculating other important evaluation methods such as accuracy, recall, precision, etc. A confusion matrix consists of four values, true positive, false positive, false negative, and true negative. The model is capable of correctly classifying all three categories even though the dataset is slightly imbalanced (**Figure 5**).

The confusion matrix was further used to calculate the precision and recall of the highest performing model. Recall is a performance metric used in classification that measures the proportion of true positive predictions out of all actual positive samples. It is also known as sensitivity or true positive rate. In the case of detecting tumors in medical images, it can be interpreted as the ability of the model to correctly identify all tumors present in the images. Recall is often considered an important metric when choosing the best model for classification of tumors. Considering that detection of both Non-Viable Tumor and Viable Tumor classes denote presence of cancer, the calculated precision for model was 1.0 and recall for the model was 0.9411.

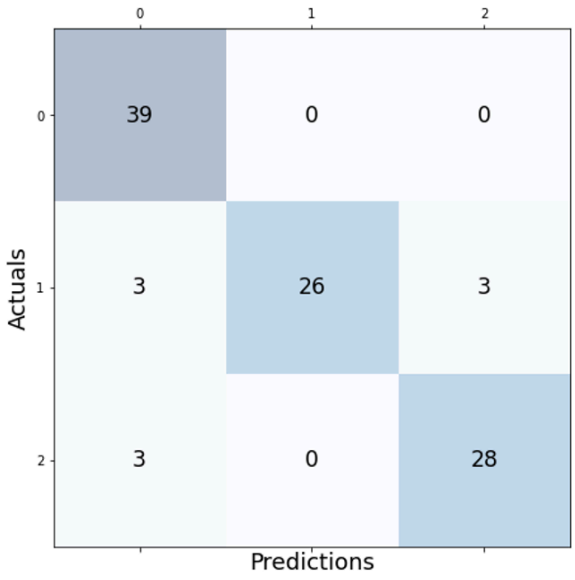

Figure 5: The confusion matrix for the highest obtained accuracy for the MobileNetV2 model. The correctly and incorrectly predicted number of images for each class for the highest reached accuracy with MobileNetV2 is shown. 0 = Non-Tumor, 1 = Non-Viable Tumor, and 2 = Viable Tumor.

DISCUSSION

Many papers using transfer learning for classifying different types of cancer have been published in recent years (10). Chang, *et al*. previously introduced a method that utilizes InceptionNetV3 to classify breast cancer histological images with transfer learning (11). The model was pre-trained on non-medical images for this task. Another paper involved using a pre-trained AlexNet and VGG-16 model instead of InceptionNetV3 and obtained high accuracy in detecting breast cancer through histologic images (12). A double transfer learning method for extracting features from histopathological images using an InceptionNetV3 convolutional neural network (CNN) and filtering the features using a trained support machine was also introduced in another paper (13).

We were able to utilize transfer learning and data augmentation for detecting osteosarcoma by classifying histological images. Our results confirmed our hypothesis that a combination of these techniques can improve detection accuracy over the reported in prior work (86% vs our 91%) (9). However, evaluating the model based on the recall value indicated more false negatives than false positives, which was an area that showed room for improvement, since false negatives risk the cancer being undetected.

We made several additional observations during our study. ResNet50V2 has 25.6 million parameters while MobileNetV2 has only 3.5 million parameters. Considering this, it was interesting to observe the differences between these two models when used for transfer learning. The results we obtained are consistent with the expectations and goals of these two networks - MobileNetV2 is targeted to achieve good accuracy while optimizing compute efficiency (rendering it suitable for a wide range of deployments including edge devices) - while ResNet50V2 is targeted for peak predictive performance. Our results suggest that both are of high quality but ResNet50V2 delivered slightly higher accuracies. It is worth noting, however, that the accuracies of both were very close - which may render MobileNetV2 to be the superior choice for implementation environments where CPU resources are limited.

In addition, ResNet50V2 outperformed MobileNetV2 when only transfer learning was used, but MobileNetV2 gave consistently better performance than any other model after data augmentation was added. This suggested that the larger size of ResNet50V2 was an advantage initially, but that MobileNetV2 was capable of achieving similar or better performance when suitable data expansion was possible. This was a notable observation because the number of parameters in MobileNetV2 is lower than other models. MobileNetV2 is much easier to deploy in a wide range of computational environments so this result suggested that our approach for detecting osteosarcoma can be widely deployed by leveraging MobileNetV2.

The small dataset available for the task was challenging but was tackled by using data augmentation through geometric

transformations. With data augmentation, using MobileNetV2 as the base model for our transfer learning, we were able to outperform the best accuracy obtained by Ahmed, *et al*., which was 86% for testing accuracy using a regularized custom CNN and a balanced set of images that utilized oversampling (9). Hence, our results suggest that existing, readily available, state-of-the-art networks like MobileNetV2 can be leveraged to obtain higher accuracy, as long as they are paired with techniques such as transfer learning and data augmentation. Given the ease of deploying MobileNetV2 (it is available on multiple platforms and is energy efficient), this can help create a broader use of neural networks for osteosarcoma detection.

We observed several limitations in the study. First, the model was biased toward predicting certain classes more accurately due to having more data points in the specific category. In the confusion matrix for predictions on the test dataset, we observed that the model was slightly biased towards predicting non-tumor images (**Figure 5**). The six false negative cases can be crucial in life threatening situations. False negative cases are dangerous since the cancer remains undiscovered, meaning that patients do not receive proper treatment. This could have been due to the imbalance present in the dataset. We also saw that overfitting started after epoch 25 because the training accuracy kept rising while the validation accuracy started falling (**Figure 4**). The overfitting issue was also observed in (9) and is an ongoing concern to be kept in mind when evaluating the results. Another limitation was runtime; since we only used basic GPUs during experimentation, the training was limited due to runtime, the time to train the models and run experiments.

In this work, we utilized several types of data augmentation. The best results were achieved when all of the data augmentation transformations used (random rotate, horizontal flip, and zoom) were combined (**Table 1**). Random rotation was tested up to 30 degrees, which reached the highest accuracies. Adding horizontal flip and zooming by a wide range of values (1.5 was the highest) also improved the accuracy. Since medical images are challenging to obtain (due to privacy concerns, lack of suitable equipment for tracking and storage), any technique to synthetically augment data in software is practically valuable.

There are several methods possible that may help improve accuracy further. It is possible that additional data augmentation could accomplish this. Other complementary augmentation techniques, such as Generative Adversarial Networks (GANs), have shown promise in other medical applications. In the future, we can extend this work by using different data augmentation methods such as generating new images using GANs. One of the main advantages of GANs is their ability to generate new, synthetic images that can be used to augment the training dataset. This can be particularly useful in medical imaging, where there is a shortage of labeled data. By generating new images, GANs can help to increase the size and diversity of the training dataset, which can lead to better-performing models. Another possible area of improvement is to change the architecture following the base model for transfer learning, which could produce higher accuracies with experimentation. We also plan to explore this in future work.

Availability of machine learning models like the one developed in this study enables automation. Such automation helps pathologists assess the tumors from histopathology images, which is a very complex task in the case of osteosarcoma (23). Developing lightweight models such as MobileNetV2 enables greater portability and usability on devices without high end computations such as GPUs.

MATERIALS AND METHODS

All experiments were run with GPUs on Google Colab software, using python libraries of TensorFlow and Keras. The dataset used in this paper is from the Children's Medical Center in Dallas, which provided a set of osteosarcoma tumor and non-tumor samples from 50 patients (19). The data was collected between 1995 and 2015 and utilized 1024x1024 labeled pixel images at 10X resolution. The distribution of the images are as follows: 536 (47%) non-tumor images, 263 (23%) necrotic tumor images and 345 (30%) viable tumor images (**Figure 2**). The images were first sorted into the correct classification before we randomly divided the complete dataset into training and validation datasets with 90% of images in the training dataset and the rest in the validation dataset. The dataset was uploaded to Google Drive, where a file path was used for the models to reference and train on.

We first tested out various TensorFlow models for transfer learning (MobileNetV2, ResNet50V2, Xception, and InceptionResNetV2 pretrained on ImageNet) and then chose two of the best performing and most consistent models to apply data augmentation. These models are all convolutional neural networks used for image classification and object detection designed to achieve high accuracies while being computationally efficient.

Different accuracies were reached by the combination of each model with different learning rates and epochs (**Figure 3**). These chosen models were trained for 30, 50 and 100 epochs using a learning rate of 0.1, 0.01, 0.005, 0.001, 0.0005, and 0.0001 with Adam optimizer (24). We used default values for beta1, beta2, and epsilon of Adam optimizer provided by TensorFlow2. The loss chosen was categorical cross entropy. The batch size selected was 32 and images were resized to 224 x 224 x 3. The training for all the trainable layers of the base model was set to false. A dense layer of size 100 and global average pooling layer were added following the base model. The output layer consisted of a softmax activation function. The accuracies were averaged after five runs at each learning rate. A plot of validation and training accuracy at each epoch helped visualize the model throughout its training process.

Finally, various types of data augmentation were added to the data as a step before training the transfer learning models with an addition to the original hyperparameters, and the best models of MobileNetV2 and ResNet50V2 were tested. First, random rotation was tested, then horizontal flip and zoom were added as parameters to the ImageDataGenerator function. Finally, the models were evaluated based on accuracy.

Received: December 03, 2022
Accepted: February 14, 2023
Published: June 03, 2023

REFERENCES

1. "Osteosarcoma - Symptoms and Treatment." *Familydoctor.org*, 6 Oct. 2020, familydoctor.org/condition/osteosarcoma/. Accessed 16 Nov. 2022.

2. Chou, Alexander J, *et al*. "Therapy for osteosarcoma: where do we go from here?." *Pediatric Drugs*, vol. 10, no. 5, 2008, pp. 315-327., doi:10.2165/00148581-200810050-00005.
3. "American Cancer Society: Cancer Facts & Statistics." *American Cancer Society | Cancer Facts & Statistics*, cancerstatisticscenter.cancer.org/#!/childhood-cancer.
4. "Key Statistics for Osteosarcoma." *American Cancer Society*, 2021, www.cancer.org/cancer/osteosarcoma/about/key-statistics.html.
5. Wittig, James C., *et al*. "Osteosarcoma: A Multidisciplinary Approach to Diagnosis and Treatment." *American Family Physician*, vol. 65, no. 6, 15 Mar. 2002, pp. 1123-1132.
6. Reason, J. "Understanding Adverse Events: Human Factors." *Quality in Health Care*, vol. 4, no. 2, 1995, pp. 80-89. doi:10.1136/qshc.4.2.80.
7. Torrey, Lisa, and Jude Shavlik. "Transfer Learning." *Handbook of Research on Machine Learning Applications and Trends: Algorithms, Methods, and Techniques*, edited by Olivas E. *et al*., IGI Global, 2010, pp. 242-264., doi:10.4018/978-1-60566-766-9.ch011.
8. Shorten, Connor, and Taghi M. Khoshgoftaar. "A Survey on Image Data Augmentation for Deep Learning." *Journal of Big Data*, vol. 6, no. 1, 2019, pp. 1-48., doi:10.1186/s40537-019-0197-0.
9. Ahmed, Imran, *et al*. "Convolutional Neural Network for Histopathological Osteosarcoma Image Classification." *CMC-Computers, Materials & Continua*, vol. 69, no. 3, 2021, pp. 3365-3381., doi:10.32604/cmc.2021.018486.
10. Yu, Xiang, *et al*. "Transfer Learning for Medical Images Analyses: A Survey." *Neurocomputing*, vol. 489, 2022, pp. 230-254., doi:10.1016/j.neucom.2021.08.159.
11. Chang, Jongwon, *et al*. "A Method for Classifying Medical Images Using Transfer Learning: A Pilot Study on Histopathology of Breast Cancer." *2017 IEEE 19th International Conference on e-Health Networking, Applications and Services (Healthcom)*, IEEE, 2017., doi:10.1109/healthcom.2017.8210843.
12. Deniz, Erkan, *et al*. "Transfer Learning Based Histopathologic Image Classification for Breast Cancer Detection." *Health Information Science and Systems*, vol. 6, no. 1, 2018, pp. 1-7., doi:10.1007/s13755-018-0057-x.
13. De Matos, Jonathan de, *et al*. "Double Transfer Learning for Breast Cancer Histopathologic Image Classification." *2019 International Joint Conference on Neural Networks (IJCNN)*, July 2019, doi:10.1109/ijcnn.2019.8852092.
14. Park, Daniel S., *et al*. "SpecAugment: A Simple Data Augmentation Method for Automatic Speech Recognition." Interspeech 2019, Sept. 2019, doi:10.21437/interspeech.2019-2680.
15. Wei, Jason, and Kai Zou. "EDA: Easy Data Augmentation Techniques for Boosting Performance on Text Classification Tasks." *Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing (EMNLP-IJCNLP)*, 2019, doi:10.18653/v1/d19-1670.
16. Shorten, Connor, *et al*. "Text Data Augmentation for Deep Learning." *Journal of Big Data*, vol. 8, no. 1, 2021, doi:10.1186/s40537-021-00492-0.
17. Luo, Yun, and Bao-Liang Lu. "EEG Data Augmentation for Emotion Recognition Using a Conditional Wasserstein GAN." *2018 40th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, 2018, doi:10.1109/embc.2018.8512865.
18. Ying, Xue. "An Overview of Overfitting and Its Solutions." *Journal of Physics: Conference Series*, vol. 1168, 2019, p. 022022. doi: 10.1088/1742-6596/1168/2/022022.
19. Leavey, P., *et al*. "Osteosarcoma data from UT Southwestern/UT Dallas for viable and necrotic tumor assessment [Data set]." *The Cancer Imaging Archive*, 2019, pp. 14.
20. Sandler, Mark, *et al*. "MobileNetV2: Inverted Residuals and Linear Bottlenecks." *2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition,* IEEE, 2018, doi:10.1109/cvpr.2018.00474.
21. Targ, Sasha, *et al*. "Resnet in resnet: Generalizing residual architectures." *arXiv preprint arXiv:1603.08029* (2016).
22. Chollet, François. "Xception: Deep Learning with Depthwise Separable Convolutions." 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), July 2017, doi:10.1109/cvpr.2017.195.
23. Mahore, Sanket, *et al*. "Machine Learning Approach to Classify and Predict Different Osteosarcoma Types." *2021 8th International Conference on Signal Processing and Integrated Networks (SPIN)*, Aug. 2021, doi: 10.1109/SPIN52536.2021.9566061.
24. Kingma, Diederik P. and Jimmy Ba. "Adam: A Method for Stochastic Optimization." *International Conference on Learning Representations (ICLR)*, 2015.