Improving Bone Fracture Detection using GANs
By Raehash Shah and Ethan Vertal

**Introduction**

In the United States, about 6.8 million people are affected by a bone fracture each year. Often an average individual is expected to sustain two fractures in their lifetime. Currently, the method of practice to detect these bone fractures are done by performing an X-ray where bone fractures are detected. Recently there has been a stronger push towards using deep learning techniques to detect these fractures. Implementation of deep learning speeds up and is more efficient in diagnosing bone fractures (Meena, 2022). Therefore developing tools that can be used to identify bone fractures in X-ray scans can be a crucial advancement in technology for the patients that are affected by these injuries which can potentially surpass human performance in detecting and classifying bone fractures from X-ray scans.

Currently there are many different deep learning models used for classification of medical imaging data. Since the introduction of CNNs in performing classification, there has been a push for using more models on clinical medical data. However, using deep learning models can have some positive consequences in highlighting specific aspects of X-ray images. By introducing models that can alter pixel intensities and perform resolution and enhancement modifications to the image can highlight specific features of X-ray images which can be used better for classification. This is where the introduction of deeper networks like ResNet and VGGNet have shown promising results. In addition, deep networks that have skipped connections like DenseNet and UNet have shown better performance since there is a boost to discriminating ability (Rahamaniar, 2019).

Since diagnosing and detection of bone imaging data requires a notion of an attention mechanism to detect “where” and “what” to concentrate on, adversarial learning would be a valuable tool to focus on. In this one can perform squeeze and excitation to control channel attention and thus improve bone fracture detection. Using this notion of attention in combination with an adversarial network like a GAN could be a potentially strong tool used in bone fracture detection (Sarvamangala, 2022).

GANs (generative-adversarial networks) are composed of a generator and a discriminator. The generator can learn the important features of the dataset and generates feature maps of the data into a latent space. Then the discriminator can distinguish these feature maps into a data distribution which we can use to identify patterns in the images like a fracture on the bone.

Therefore over the course of this project, the goal will be to create such a deep learning model that can outperform some of the current models used as a part of X-ray detection of bone fractures. As an input, we will work with a dataset that is composed of fractures and non-fractured X-ray images of multiple different joints that are found in the upper-extremities. We will baseline our model with a classical imaging classification algorithm like a CNN. Then we will show improvements to our classification of our data by using deeper networks like a

UNet and then working with a GAN that show even further improvement in our classification accuracy.

**Preliminary Results**

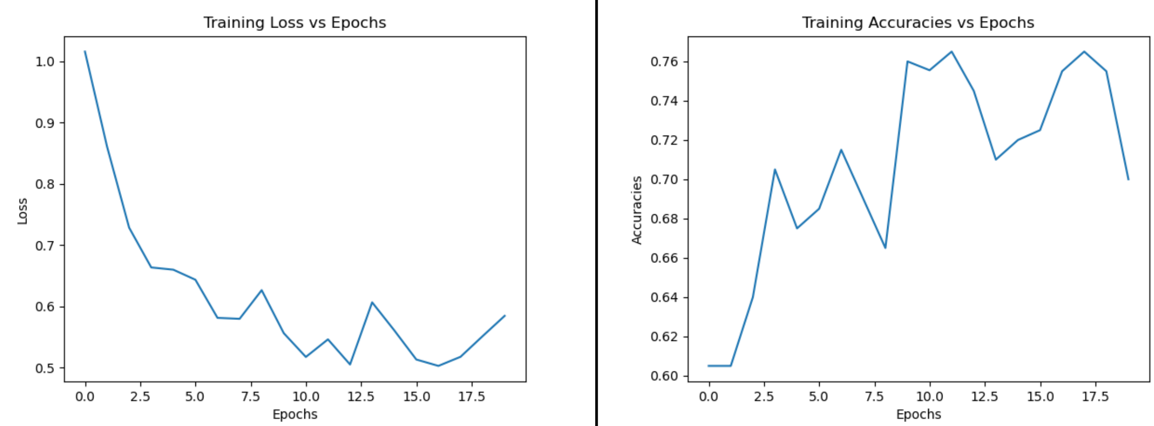


***Figure 1: Training Loss and Accuracy for the CNN Baseline Model***
Notice that the general trend in the training loss graph is decreasing suggesting the model is minimizing loss and learning. Similarly, the model's accuracy in classifying is improving suggesting learning as well.

As an initial attempt at producing some successful results, we developed a type of CNN. Using inspiration from previous existing image based CNN models, our model referenced architecture from the EfficientNet (Elmoufidi, 2021) which had some promising results on ImageNet and a Diabetic Retinopathy dataset. In this model, we have multiple convolutional layers with IRC which are inverted residual connections. We made slight modifications by adding gaussian noise and dropout layers to reach a better performance on the model. We split our data into a training set composed of 8863 images and a testing set of 600 images each with the two classes of whether a fracture is present or not. After training it on 20 epochs with 50 steps in each epoch, results shown above, the model reached a training accuracy of 75.31% and a testing accuracy of 65.90%.

While our model produced lower accuracy than the original EfficientNet model did on their data set, these results are promising. The inclusion of residual connections and the additional noise and dropout layers here helped us determine that they will prove useful in developing our UNet architecture as we move forward with model experimentation. We will be expanding on a UNet due to its proven capability of segmenting biomedical images (Ronneberger et al., 2015). We will be experimenting with, not only just the hyperparameters and shape of the "U" used previously in their models, but also the type of feature engineering we can perform on the data prior to passing it into the network. By creating different stylized features out of our data, we could find greater improvements in accuracy in fracture detection and classification.

Though 75% is a good start, this number is nowhere near the goal of performing better than a trained doctor's eye. In order to step above this, we additionally plan on implementing a GAN, as discussed above, to attempt to better classify the images. It has been shown that if we extract a GAN's discriminator, we can actually classify images with comparable or better results than using a regular CNN. Further, if we were to use convolutions in our discriminator network, we can expect better results than a CNN in image classification (Meng and Guo, 2021). If these results are strong enough, further experiments would be nice if we have time but are not necessarily part of our main question of whether a GAN does better at detecting and classifying bone fractures. However, if we are not satisfied, we may look into adding our residual connections, dropout layers, and noise layers into the discriminator to enhance its performance.

**Tentative Timeline:**

| Task: | Description: | Due By: | Member(s): |
|---|---|---|---|
| UNet & GAN architectures | Develop a preliminary structure to experiment with for each type of model. | 11/06 - Monday | Both |
| UNet Finalized | Finish UNet training, optimization<br><br>Should wrap up when satisfied with results (accuracy no lower than 95%) | 11/20 or 27 - Mondays, depending on work load pre-Thanksgiving | Ethan |
| GAN Finalized | Finish GAN training, optimization<br><br>Should wrap up when satisfied with results (accuracy no lower than 95%) | 11/27 - Monday | Raehash |
| Documentation | Document, report, and summarize findings along the way to save work for us down the line | As we go | Both |
| Final Paper | Write final report | 12/1 - Friday | Both |

**References:**

Elmoufidi, Abdelali and H Bazm Amoun. “EfficientNetB3 Architecture for Diabetic Retinopathy Assessment using Fundus Images.” (2021).

Meena, Tanushree, and Sudipta Roy. “Bone Fracture Detection Using Deep Supervised Learning from Radiological Images: A Paradigm Shift.” *Diagnostics (Basel, Switzerland)* vol. 12,10 2420. 7 Oct. 2022, doi:10.3390/diagnostics12102420

Rahmaniar, Wahyu, and Wen-June Wang. "Real-time automated segmentation and classification of calcaneal fractures in CT images." *Applied Sciences* 9.15 (2019): 3011.

Ronneberger, Olaf, et al. “U-Net: Convolutional Networks for Biomedical Image Segmentation.” *Arxiv*, Cornell University, 2015, arxiv.org/pdf/1505.04597.pdf.

Sarvamangala, D. R., and Raghavendra V. Kulkarni. "Convolutional neural networks in medical image understanding: a survey." *Evolutionary intelligence* 15.1 (2022): 1-22.

H. Meng and F. Guo, "Image Classification and Generation Based on GAN Model," 2021 3rd International Conference on Machine Learning, Big Data and Business Intelligence (MLBDBI), Taiyuan, China, 2021, pp. 180-183, doi: 10.1109/MLBDBI54094.2021.00042.