



Osteosarcoma data from UT Southwestern/UT Dallas for Viable and Necrotic Tumor Assessment

Summary

Osteosarcoma is the most common type of bone cancer that occurs in adolescents in the age of 10 to 14 years. The dataset is composed of Hematoxylin and eosin (H&E) stained osteosarcoma histology images. The data was collected by a team of clinical scientists at University of Texas Southwestern Medical Center, Dallas. Archival samples for 50 patients treated at Children’s Medical Center, Dallas, between 1995 and 2015, were used to create this dataset. Four patients (out of 50) were selected by pathologists based on diversity of tumor specimens after surgical resection. The images are labelled as Non-Tumor, Viable Tumor and Necrosis according to the predominant cancer type in each image. The annotation was performed by two medical experts. All images were divided between two pathologists for the annotation activity. Each image had a single annotation as any given image was annotated by only one pathologist. The dataset consists of 1144 images of size 1024 X 1024 at 10X resolution with the following distribution: 536 (47%) non-tumor images, 263 (23%) necrotic tumor images and 345 (30%) viable tumor tiles.

Data Access

Data Access

Data Type	Download all or Query/Filter
Images (JPG, 196MB)	
Features (CSV)	

Click the Versions tab for more info about data releases.

Detailed Description

Detailed Description

Image Statistics	
Modalities	Pathology
Number of Participants	4
Number of Images	1144
Images Size (MB)	196

Folder_Structure

- Data_Osteo_Files
 - ML_Features_1144.csv - Contains 1144 rows for all the image tiles and 69 columns for filename, classification, and 65 machine learning features.

- Training_Set_1 - 11 folders with 547 images. Each folder contains 48~50 image tiles and 1 csv for annotation.
 - set 1- 49 Image Tiles
 - set 2- 50 Image Tiles
 - set 3- 50 Image Tiles
 - set 4- 50 Image Tiles
 - set 5- 50 Image Tiles
 - set 6- 50 Image Tiles
 - set 7- 50 Image Tiles
 - set 8- 50 Image Tiles
 - set 9- 50 Image Tiles
 - set 10- 50 Image Tiles
 - set 11- 48 Image Tiles
- Training_Set_2 - 12 folders with 597 images. Each folder contains 48~50 image tiles and 1 csv for annotation.
 - set 1- 49 Image Tiles
 - set 2- 50 Image Tiles
 - set 3- 50 Image Tiles
 - set 4- 50 Image Tiles
 - set 5- 50 Image Tiles
 - set 6- 50 Image Tiles
 - set 7- 50 Image Tiles
 - set 8- 50 Image Tiles
 - set 9- 50 Image Tiles
 - set 10- 50 Image Tiles
 - set 11- 50 Image Tiles
 - set 12- 48 Image Tiles

Citations & Data Usage Policy

Citations & Data Usage Policy

Users of this data must abide by the [TCIA Data Usage Policy](#) and the [Creative Commons Attribution 3.0 Unported License](#) under which it has been published.

Attribution should include references to the following citations:

Data Citation

Leavey, P., Sengupta, A., Rakheja, D., Daescu, O., Arunachalam, H. B., & Mishra, R. (2019). Osteosarcoma data from UT Southwestern /UT Dallas for Viable and Necrotic Tumor Assessment [Data set]. The Cancer Imaging Archive. <https://doi.org/10.7937/tcia.2019.bvhjhdas>

Publication Citation

1) Mishra, Rashika, et al. "Histopathological diagnosis for viable and non-viable tumor prediction for osteosarcoma using convolutional neural network." International Symposium on Bioinformatics Research and Applications. Springer, Cham, 2017.

2) Arunachalam, Harish Babu, et al. "Computer aided image segmentation and classification for viable and non-viable tumor identification in osteosarcoma." PACIFIC SYMPOSIUM ON BIOCOMPUTING 2017. 2017.

3) Mishra, Rashika, et al. "Convolutional Neural Network for Histopathological Analysis of Osteosarcoma." Journal of Computational Biology 25.3 (2018): 313-325.

4) Leavey, Patrick, et al. "Implementation of Computer-Based Image Pattern Recognition Algorithms to Interpret Tumor Necrosis; a First Step in Development of a Novel Biomarker in Osteosarcoma." PEDIATRIC BLOOD & CANCER. Vol. 64. 111 RIVER ST, HOBOKEN 07030-5774, NJ USA: WILEY, 2017.

TCIA Citation


Clark K, Vendt B, Smith K, Freymann J, Kirby J, Koppel P, Moore S, Phillips S, Maffitt D, Pringle M, Tarbox L, Prior F. **The Cancer Imaging Archive (TCIA): Maintaining and Operating a Public Information Repository**, Journal of Digital Imaging, Volume 26, Number 6, December, 2013, pp 1045-1057. DOI: [10.1007/s10278-013-9622-7](https://doi.org/10.1007/s10278-013-9622-7)

Other Publications Using This Data

TCIA maintains [a list of publications](#) which leverage TCIA data. If you have a manuscript you'd like to add please [contact the TCIA Helpdesk](#).

Versions

Version 1 (Current): Updated 2019/03/22

Data Type	Download all or Query/Filter
Images (JPG, 196MB)	
Features (CSV)	