



# **SN-AM Dataset: White Blood cancer dataset of B-ALL and MM for stain normalization**


## Summary

Microscopic images were captured from bone marrow aspirate slides of patients diagnosed with B-lineage Acute Lymphoid Leukemia (B-ALL) and Multiple Myeloma (MM) as per the standard guidelines. Slides were stained using Jenner-Giemsa stain. Images were captured at 1000x magnification using Nikon Eclipse-200 microscope equipped with a digital camera. Images were captured in raw BMP format with a size of 2560x1920 pixels. In all, this dataset consists of 90 images of B-ALL and 100 images of MM. Both MM and B-ALL images have sufficient variability from one image to another image to rigorously test any stain normalization methodology developed. More information about each subset are provided on the Detailed Description tab below.

### Data Access

#### Data Access

Click the **Download** button to browse and download the data from Box.

Data Type	Download all or Query/Filter
Images (BMP, TIFF, 2.9 GB)	

Click the Versions tab for more info about data releases.

### Detailed Description

#### Detailed Description

Image Statistics	
Modalities	Pathology
Number of Participants	16
Number of Studies	60
Number of Images	190
Images Size (GB)	2.9

#### Data subset-1: ALL images

Microscopic images were captured from bone marrow aspirate slides of patients diagnosed with B-lineage Acute Lymphoblastic Leukemia (B-ALL). Slides were stained using Jenner-Giemsa stain and lymphoblasts, that are cells of interest, have been evaluated. Images were captured in raw BMP format with a size of 2560x1920 pixels using Nikon Eclipse-200 microscope equipped with a digital camera at 1000x magnification. In all, this dataset consists of 30 images, wherein one image has been used as the reference image and the proposed stain normalization method has been tested on 29 images. For each of these 30 images, we have also provided two additional images that contain the nucleus mask and the background mask, respectively, for that particular image. For example, if the original file is saved with the name “ALL\_1.bmp”, the corresponding image with mask on the nuclei is saved as “ALL\_1\_nucleus\_mask.bmp”, and the corresponding image with mask on the background is saved as “ALL\_1\_background\_mask.bmp. Thus, in all, we have 90 images for this dataset.

### **Data subset-2: MM images**

The third data subset contains microscopic images captured from slides prepared from bone marrow aspirate collected from patients with Multiple Myeloma (MM). Slides are stained using Jenner-Giemsa stain and plasma cells, that are cells of interest, have been evaluated. A total of 30 images have been considered, wherein one image has been used as the reference image to which 29 images have been stain normalized. For each of these 30 images, we have also provided two additional images that contain the nucleus mask and the background mask, respectively, for that particular image. For example, if the original file is saved with the name “MM\_1.bmp”, the corresponding image with mask on the nuclei is saved as “MM\_1\_nucleus\_mask.bmp”, and the corresponding image with mask on the background is saved as “MM\_1\_background\_mask.bmp. In addition, for 17 images, the mask images are also provided for the cytoplasm of the plasma cells, namely, “MM\_1\_cyto\_mask.bmp. Thus, in all, we have 100 images for this dataset.

### **Citations & Data Usage Policy**

## **Citations & Data Usage Policy**

These collections are freely available to browse, download, and use for commercial, scientific and educational purposes as outlined in the [Creative Commons Attribution 3.0 Unported License](#). Questions may be directed to [help@canerimagingarchive.net](mailto:help@canerimagingarchive.net). Please be sure to acknowledge both this data set and TCIA in publications by including the following citations in your work:

#### **Data Citation**

"Gupta, A., & Gupta, R. (2019). SN-AM Dataset: White Blood Cancer Dataset of B-ALL and MM for Stain Normalization [Data set]. The Cancer Imaging Archive. <https://doi.org/10.7937/tcia.2019.of2w81xr>"

### Publication Citation

1. Anubha Gupta, Rahul Duggal, Ritu Gupta, Lalit Kumar, Nisarg Thakkar, and Devprakash Satpathy, "GCTI-SN: Geometry-Inspired Chemical and Tissue Invariant Stain Normalization of Microscopic Medical Images", under review.
2. Ritu Gupta, Pramit Mallick, Rahul Duggal, Anubha Gupta, and Ojaswa Sharma, "Stain Color Normalization and Segmentation of Plasma Cells in Microscopic Images as a Prelude to Development of Computer Assisted Automated Disease Diagnostic Tool in Multiple Myeloma," 16th International Myeloma Workshop (IMW), India, March 2017.
3. Rahul Duggal, Anubha Gupta, Ritu Gupta, and Pramit Mallick, "SD-Layer: Stain Deconvolutional Layer for CNNs in Medical Microscopic Imaging," In: Descoteaux M., Maier-Hein L., Franz A., Jannin P., Collins D., Duchesne S. (eds) Medical Image Computing and Computer-Assisted Intervention MICCAI 2017, MICCAI 2017. Lecture Notes in Computer Science, Part III, LNCS 10435, pp. 435–443. Springer, Cham. DOI: [https://doi.org/10.1007/978-3-319-66179-7\\_50](https://doi.org/10.1007/978-3-319-66179-7_50).
4. Rahul Duggal, Anubha Gupta, Ritu Gupta, Manya Wadhwa, and Chirag Ahuja, "Overlapping Cell Nuclei Segmentation in Microscopic Images Using Deep Belief Networks," Indian Conference on Computer Vision, Graphics and Image Processing (ICVGIP), India, December 2016.
5. Rahul Duggal, Anubha Gupta, and Ritu Gupta, "Segmentation of overlapping/touching white blood cell nuclei using artificial neural networks," CME Series on Hemato-Oncopathology, All India Institute of Medical Sciences (AIIMS), New Delhi, India, July 2016.

### 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 2019/03/26

Data Type	Download all or Query/Filter
-----------	------------------------------

Images (BMP, TIFF, 2.9 GB)	 A blue rectangular button with rounded corners. On the left side, there is a white downward-pointing arrow icon. To the right of the icon, the word "Download" is written in white, sans-serif font.
----------------------------	--