ELSEVIER

Blood 142 (2023) 2841–2842

The 65th ASH Annual Meeting Abstracts

POSTER ABSTRACTS

613.ACUTE MYELOID LEUKEMIAS: CLINICAL AND EPIDEMIOLOGICAL

Deep Learning for Morphology-Based, Bone Marrow Cell Classification

Shenghuan Sun[1], *Jacob Cleave*[2], *Linlin Wang*[3], *Fabienne Lucas*[4], *Laura Brown*[5], *Jacob Spector*[6], *Leonardo Boiocchi*[2], *Jeeyeon Baik, BA, MPH*[7], *Menglei Zhu*[2], *Orly Ardon*[2], *Chuanyi M. Lu, MDPhD*[8], *Ahmet Dogan, MD PhD*[7], *Dmitry Goldgof*[9], *Iain Carmichael*[10], *Sonam Prakash*[5], *Atul Butte*[5], *Gregory Mark Goldgof*[11]

[1] University of California, San Francisco, San Francisco, CA
[2] Memorial Sloan Kettering Cancer Center, New York, NY
[3] UCSF, San Francisco
[4] Harvard Medical School, Boston
[5] University of California, San Francisco, San Francisco
[6] Boston Children's Hospital, Boston, MA
[7] Department of Pathology and Laboratory Medicine, Hematopathology Service, Memorial Sloan Kettering Cancer Center, New York, NY
[8] United States Department of Veterans Affairs, San Francisco, CA
[9] University of South Florida, Tampa, FL
[10] University of California, Berkeley, Berkeley, CA
[11] Memorial Sloan Kettering Cancer Center, Yew York

The morphological classification of cells in bone marrow aspirate (BMA) is central to the diagnosis of hematologic diseases, including leukemias. Despite being a critical task, its monotonous, time-consuming nature and dependency on highly skilled clinical experts makes it prone to human error. Such errors can lead to delays and misdiagnoses that negatively impact patient care. To counter these challenges, we curated an expansive dataset of more than 40,000 hematopathologist consensus-annotated single-cell images, extracted from BMA whole slide images (WSIs), each annotated into one of 23 distinct morphologic classes.

We then utilized this data to develop DeepHeme, a convolutional neural network classifier designed for bone marrow cell typing tasks. DeepHeme achieves state-of-the-art performance in both the breadth of differentiable classes and accuracy across these classes. By comparing its performance to that of individual hematopathologists from three premier academic medical centers, using our gold standard consensus-labelled images, we found our AI algorithm either matched or surpassed the average performance across all classes. In addition, we integrated DeepHeme with internally developed region classifier and cell detection algorithms, culminating in a comprehensive diagnostic pipeline for whole slide cell differential.

We next tested DeepHeme on slides from an external hospital system at a major cancer center to evaluate the generalizability of our model, a necessary precondition to widespread application. DeepHeme demonstrated a high level of generalizability, evidenced by a decrease of only 4% in the mean F-1 score, from 0.89 to 0.85, across all 23 cell classes. Lastly, to improve access to the DeepHeme algorithm results and encourage further real-world generalizability testing, we developed a web application that allows scientists and clinicians to test the DeepHeme algorithm on either test images from our study or their own user-uploaded aspirates.

Disclosures Baik: *Pauling.AI:* Current Employment. **Dogan:** *Seattle Genetics:* Consultancy; *Physicians' Education Resource:* Consultancy, Honoraria; *EUSA Pharma:* Consultancy; *Loxo:* Consultancy; *Peer View:* Honoraria; *Incyte:* Consultancy; *Takeda:* Other: Research Funding; *Roche:* Other: Research Funding.

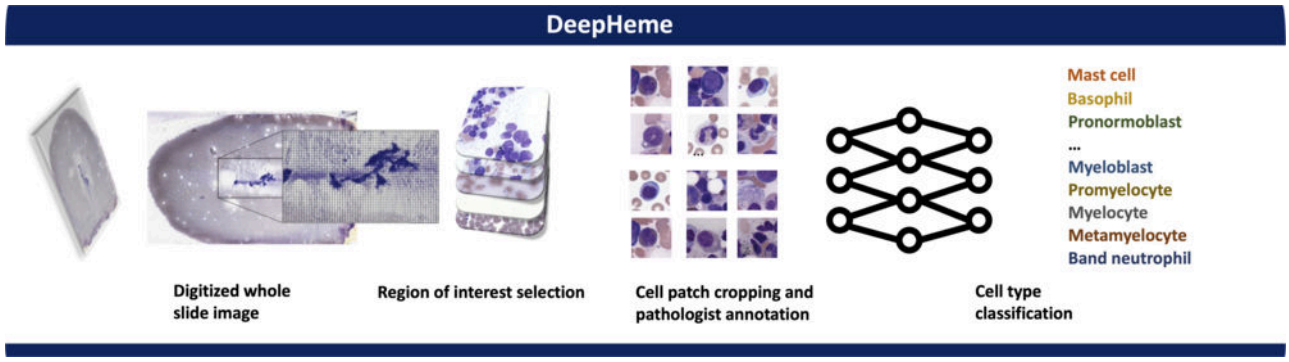

Figure 1

https://doi.org/10.1182/blood-2023-172654