**Legal Commitment for Using the** **M5HisDoc Dataset**

**Licensor:** Deep Learning and Vision Computing Lab (DLVC-Lab)

South China University of Technology (SCUT), Wushan Road, Guangzhou, China

**User: \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_­­ Institution: \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_­­\_\_\_\_\_\_**

**Address: \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_**

The Deep Learning and Vision Computing Lab created the M5HisDoc dataset for research in Chinese historical document analysis. To ensure the correct and legal use of the M5HisDoc dataset (hereinafter “Dataset”) and prevent incidents such as data leakage, the User intends to enter into this Legal Commitment. The User is willing to be legally bound and promises to be subject to the following terms of use:

1. The Licensor grants the User the right to use the Dataset, but not the ownership. The Licensor retains the copyright of all sample data in the Dataset.
2. The Dataset must be used under the [Creative Attribution-NonCommercial-NoDerivatives 4.0 International (CC BY-NC-ND 4.0) License](https://creativecommons.org/licenses/by-nc-nd/4.0/).
3. All samples in the Dataset in this commitment can only be used by the team of the named User and can only be used for academic research purposes. No samples can be used for any commercial purpose. In addition, the User shall not use the M5HisDoc dataset for purposes that are not expressly set forth in this commitment, except with the written permission of the Licensor.
4. The User shall abide by the data protection regulations stipulated by the laws of the People's Republic of China, and is required to comply with all data protection laws that may affect him/her. The User shall not use the Dataset for any illegal activity and shall bear all legal responsibilities; otherwise, the User shall be liable for all legal consequences and the authorization will be permanently revoked.
5. The User is required to list 1-2 of the publications in recent 6 years to indicate that he/she (or the team) do research in the related research fields of OCR, historical document analysis, handwriting analysis and recognition, document image processing, etc. in the application form.
6. The signature of the leader and the official seal of the institution (university, research institute, related department in the company, etc.) to which the User belongs must be provided; otherwise, the application will be rejected.
7. The User must not copy or disseminate the Dataset without authorization and no part of the Dataset may be reproduced or distributed in any form or any means.
8. All usages of the Dataset should be appropriately acknowledged. In publications, the Dataset should be referred to “M5HisDoc dataset”. Publications of research results related to the Dataset should also be appropriately acknowledged.

Signature (User): \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ Seal (User): \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

Date: \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_