

---

# Towards Visual Text Design Transfer Across Languages: Supplementary Material

---

## 1 MuST-Bench Dataset

Access to the dataset is unrestricted and does not require any authorization requests. The instructions, code, and data for the assembly of the MuST-Bench dataset are available via the following link:

[https://drive.google.com/file/d/1\\_4d0LRhhWfhTD-aNjv7RbM7xYHXZKJSM/view?usp=drive\\_link](https://drive.google.com/file/d/1_4d0LRhhWfhTD-aNjv7RbM7xYHXZKJSM/view?usp=drive_link)

## 2 MuST-Bench Metadata

To provide a comprehensive overview of the MuST-Bench dataset and its structure, below is a detailed explanation of the metadata associated with our collection of multilingual film poster images.

- Our dataset comprising multilingual film poster images was sourced from a movie database website<sup>1</sup> and serves as the raw material for our study.
- The dataset is licensed under the CC-BY-NC<sup>2</sup> license.
- The "data.json" file is organized to reflect detailed metadata for each image in the dataset. Each entry in the JSON file is structured as follows:
  - **Image ID:** Each image file, such as "1421.jpg", is uniquely identified and serves as the key in our dataset structure.
  - **Regional Variants:** Metadata for different regional variants of the film posters is categorized under country codes (e.g., US, CN, KR, RU, TH, AR). Each regional entry includes:
    - \* **Coordinates:** Character-level bounding box annotations are represented as coordinates.
    - \* **Availability:** A boolean indicates whether the poster is available.
    - \* **Link:** A direct URL to download the poster image.

## 3 Dataset Hosting and Maintenance Plan

Our dataset is not offered as raw data but rather as a structured resource including json file, data reconstruction code, and instructions. This format allows for straightforward reconfiguration and updates. The storage utilized is freely available and sufficient for the dataset's needs, ensuring that maintenance and access do not pose any challenges. The platform's robust infrastructure provides reliable and continuous access to the dataset, requiring minimal maintenance efforts.

---

<sup>1</sup><https://www.cinematerial.com/>

<sup>2</sup><https://creativecommons.org/licenses/by-nc/4.0/>

## 29 **4 Author Statement**

30 We, the authors and creators of the MuST-Bench Dataset, hereby affirm our full responsibility for the  
31 compilation and maintenance of this dataset.

## 32 **5 SIGIL Code**

33 Access to the code is unrestricted and does not require any authorization requests. The instructions,  
34 codes are available via the following link:

35 [https://drive.google.com/file/d/153R6Dtgt-WxdvF3KcXxrab0x7y-v51a0/view?usp=](https://drive.google.com/file/d/153R6Dtgt-WxdvF3KcXxrab0x7y-v51a0/view?usp=sharing)  
36 [sharing](https://drive.google.com/file/d/153R6Dtgt-WxdvF3KcXxrab0x7y-v51a0/view?usp=sharing)