Example of interleaved image-text documents:

......, .There are many different standards of analog video. Most are conceptually similar, but they have varying numbers of signal channels, ,They are all based on the analog signal(s) drawing the......

🤔 *How to utilize interleaved documents for representation learning?*
How to leverage realistic&synthetic texts to enhance representation?

Real-World Data Extraction

Vision Expert Models

Text

Image

Caption

Tag

Hierarchical Retrieval

Image Semantic Augmented Generation

Realistic Text

Synthetic Text

Text Augmentation Supervision