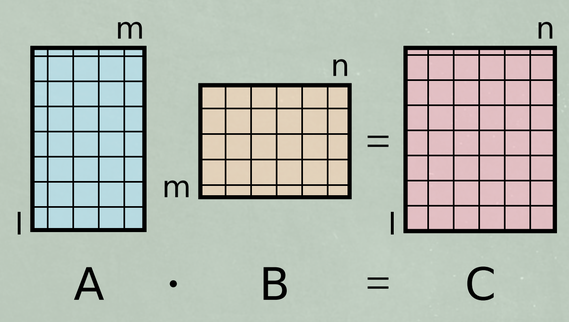
Enriching Textbooks with Images for Enhanced Learning



$$c_{ij} = a_{i1}b_{1j} + a_{i2}b_{2j} + \ldots + a_{in}b_{nj} = \sum_{k=1}^{n} a_{ik}b_{ki}$$

Matrix

In mathematics, particularly in linear algebra, matrix multiplication is a binary operation that produces a matrix from two matrices. For matrix multiplication, the number of columns in the first matrix must be equal to the number of rows in the second matrix. The resulting matrix, known as the matrix product, has the number of rows of the first and the number of columns of the second matrix. The product of matrices A and B is denoted as AB.

We use the following notational conventions: matrices are represented by capital letters in bold, e.g. **A**; vectors in lowercase bold, e.g. **a**; and entries of vectors and matrices are italic (they are numbers from a field), e.g. *A* and *a*. Index notation is often the clearest way to express definitions, and is used as standard in the literature. The entry in row i, column j of matrix **A** is indicated by $(\mathbf{A})_{ij}$, A_{ij} or a_{ij}. In contrast, a single subscript, e.g. $\mathbf{A}_1$, $\mathbf{A}_2$, is used to select a matrix (not a matrix entry) from a collection of matrices.

Good illustrations cover novel concepts

Images help learning but editing is expert labor-intensive and effort is centered on textbooks in English.

Tools for large-scale & reliable textbook enrichment help the Open Educational Resources initiative

Given an image bank, how to assign some of the images to textbook parts?

Finetuned CLIP for image-concept .
No control for multiple images covering the same concept.

maximum coverage is NP-hard
→ subset modularity
→ greedy approximation is approximately optimal

subset submodular if adding new items to $\mathcal{I}'$ always has diminishing returns

Formalization of methods for image assignment & proofs of approximations.

Dataset of textbooks with assigned images

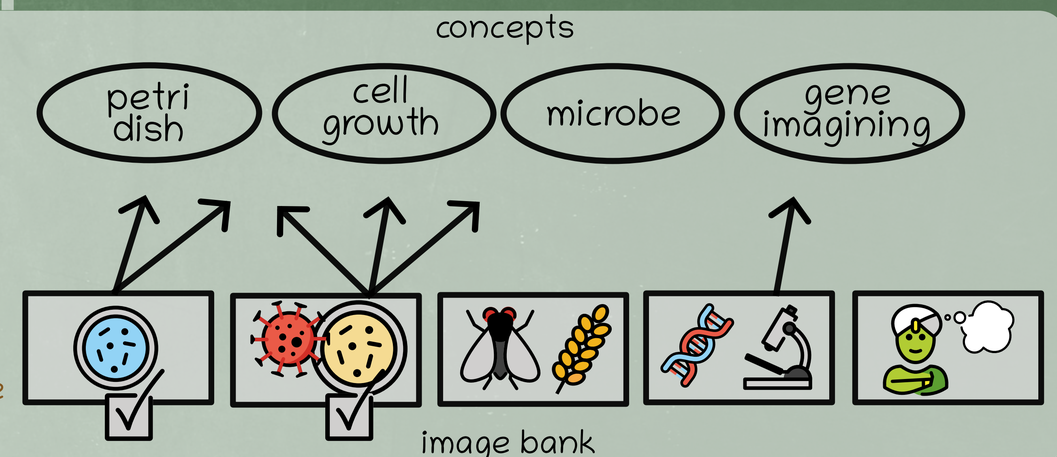
Human evaluation (teachers) of textbook image assignments

Local Method

Select top images that covers the most concepts

$$\max_{\mathcal{I}'} S(\mathcal{I}', s) = \max_{\mathcal{I}'} \sum_{i \in \mathcal{I}'} \sum_{t \in u} \text{sim}(i, t)$$

assigned images · concepts in a subsection · concept-image similarity

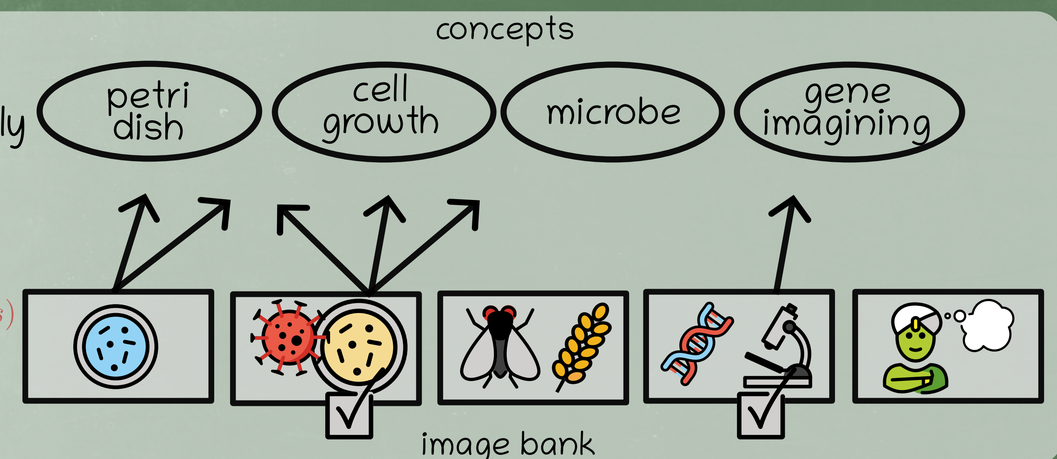


Joint Method

Select top images that jointly cover the most concepts

$$S(\mathcal{I}', s) + \beta \cdot C(\mathcal{I}', s) - \beta \cdot R(\mathcal{I}', s) = \sum_{i \in \mathcal{I}'} \sum_{t \in s} \text{sim}(i, t) + \beta \cdot C(\mathcal{I}', s) - \beta \cdot R(\mathcal{I}', s)$$

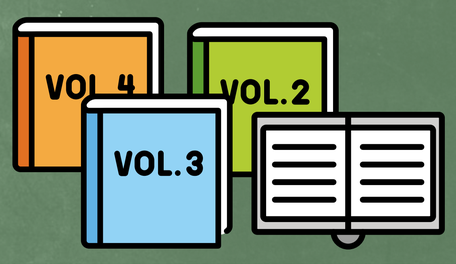
β is tradeoff between relevancy and coverage without redundancy

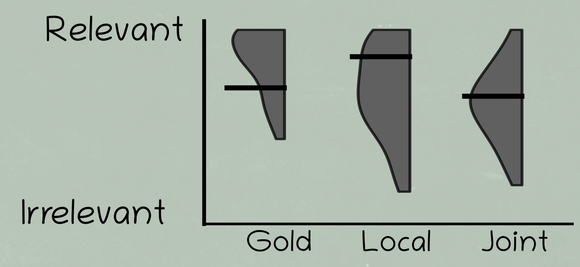


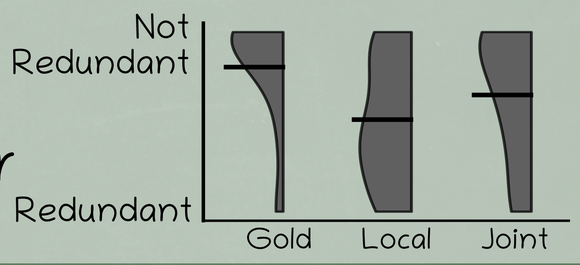
$C(\mathcal{I}', s)$
= unique covered concepts
$= \left|\{c \in s \mid \exists i \in \mathcal{I}' : \text{cov}(c, i) = 1\}\right|$

$R(\mathcal{I}', s)$
= multiply covered concepts
$= \sum_{c \in s} \sum_{i \in \mathcal{I}'} \text{cov}(c, i) - C(\mathcal{I}', s)$

35 large textbooks from maths, science, business from openstax.org with list of concepts
· 3k sections · 16k subsections · 10k assigned images
· 300k images from Wikipedia







Gap between human and automatic assignment.

Local Method has more relevant images but higher overlap. Joint Method trades the relevancy for lower redundancy.

Enhancing Textbooks with Visuals from the Web for Improved Learning
Janvijay Singh, Vilém Zouhar, Mrinmaya Sachan

ETH zürich