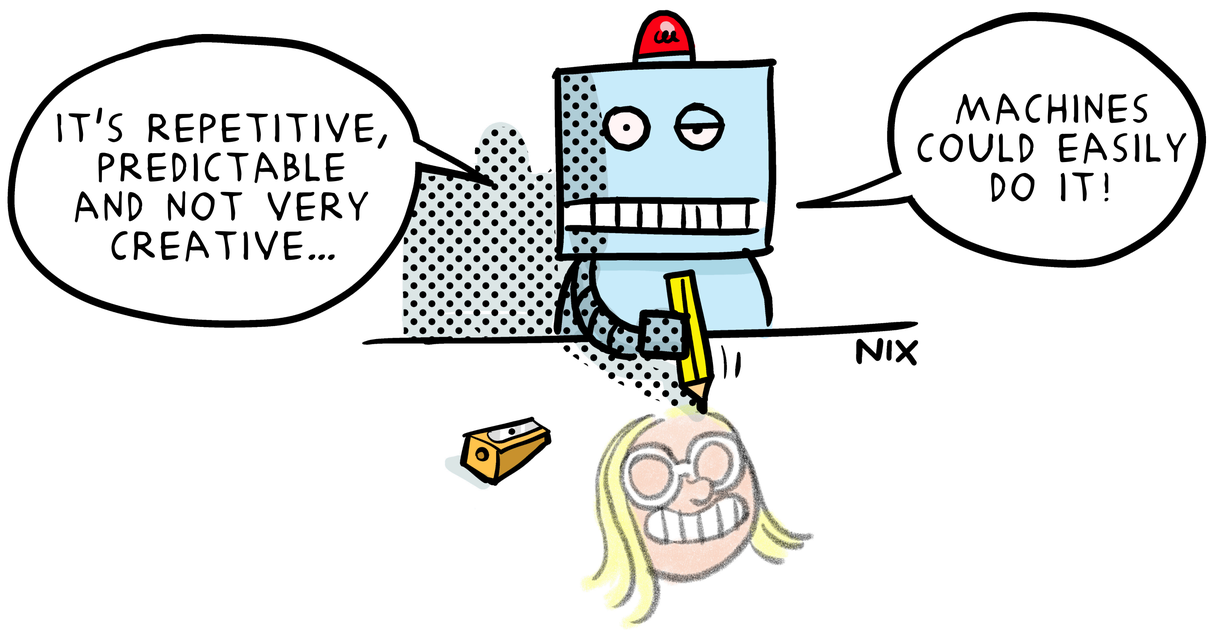
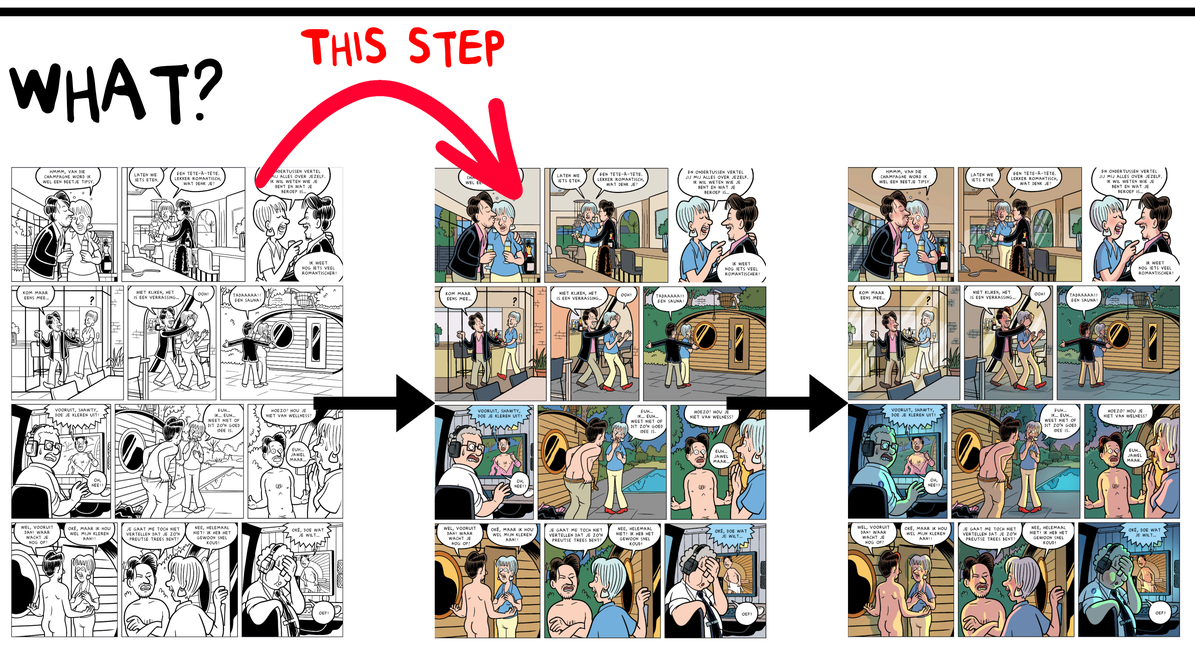
FLAT COLORING OF COMICS IS BORING!



WHAT?

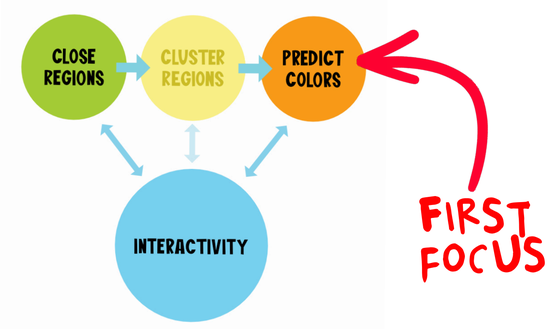


Inked page | Flat colors | Effects (shadows, ambiance)

WHY?

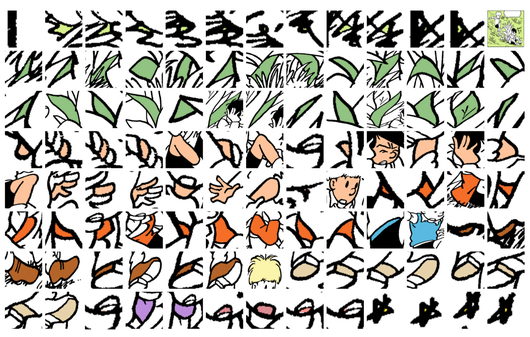
- Labor-intensive: Requires at least 2 hours per page.
- Repetitive: Ensures consistency throughout a comic book or series.
- Minimally creative: Involves primarily using the bucket tool in Photoshop or Clip Studio Paint.
- Automation: Yields time savings and makes artist's job more interesting with more time for creative tasks

The real challenge lies in a **partially** automated approach that supports the work of a **professional** artist in an **interactive** way with maximum **steerability**, without adding extra workload. Accurate color prediction is the subject of this pilot study.

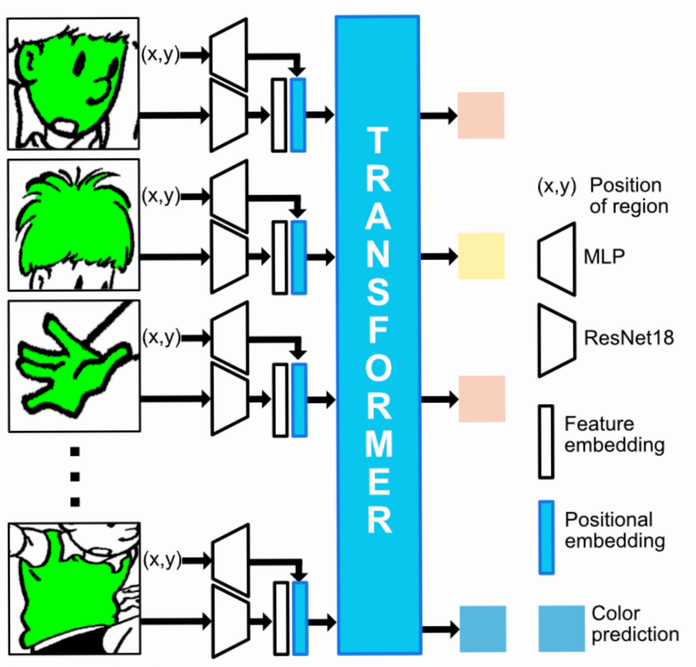


HOW?

An original panel



The extracted regions



The architecture

- Vision Transformer (ViT) processes entire panels, treating regions as tokens to focus on inter-region relationships.
- Each region tokenized using feature embeddings from an initially trained ResNet18.
- Panel centers' x, y coordinates converted via an MLP with one hidden layer.
- Sequences standardized to 200 tokens via padding.
- Model features four self-attention layers with four heads each, using cross-entropy loss for classification.

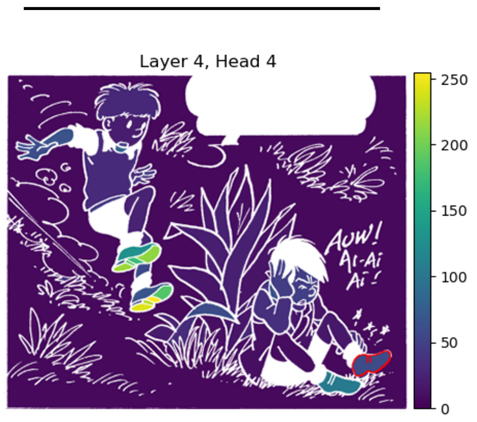
- Pilot study limited to one comic book.
- 255 panels analyzed.
- Panels segmented; regions labeled by original colors.
- Total of 37,578 regions across 248 colors.
- Regions isolated into rectangles with 20-pixel padding.
- Target colors in regions marked green (0,255,0).
- Data shuffled and split into an 80%-20% train-validation ratio.

RESULTS?

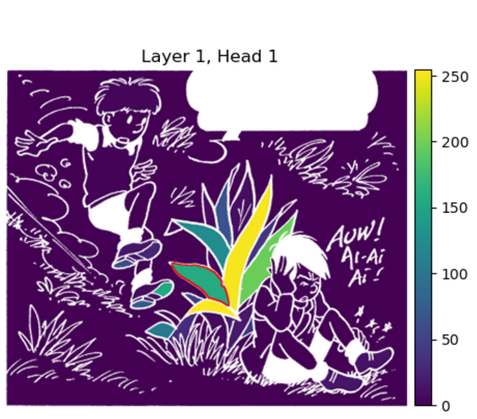
Table 1: Ablation study

ResNet	ViT PE		Acc.
yes	no	no	0.41
frozen	yes	no	0.51
frozen	yes	yes	0.54
finetuning	yes	yes	0.55

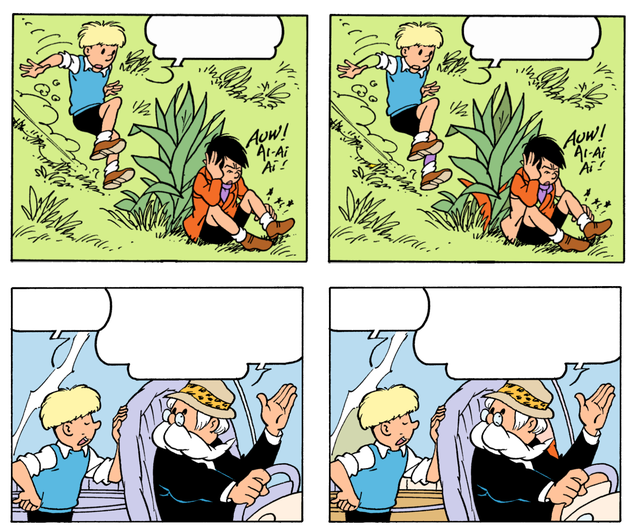
- Ablation study highlights impact of positional embeddings and ResNet18 fine-tuning.
- Attention maps reveal transformer's grasp of visual and semantic region relationships.
- Method yields satisfactory results despite modest 0.55 accuracy rate in coloring validation set panels.



Attention map for foot region



Attention map for leaf region



Original | Predicted

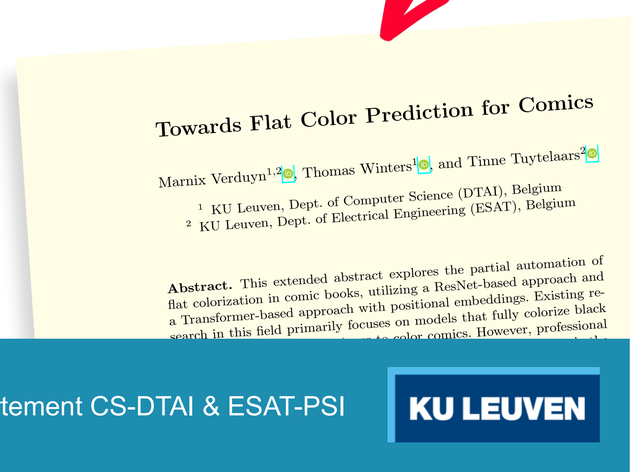
Towards Flat Color Prediction for Comics

- Extended abstract by Marnix Verduyn, Thomas Winters and Tinne Tuytelaars

Acknowledgements

- MV is supported through VLAIO Baekeland, Standaard Uitgeverij and Studio 100.
- TW received a grant from Internal Funds KU Leuven (PDMT2/23/050).

READ THE PAPER



Towards Flat Color Prediction for Comics

Marnix Verduyn[1,2], Thomas Winters[1], and Tinne Tuytelaars[2]

[1] KU Leuven, Dept. of Computer Science (DTAI), Belgium
[2] KU Leuven, Dept. of Electrical Engineering (ESAT), Belgium

Abstract. This extended abstract explores the partial automation of flat colorization in comic books, utilizing a ResNet-based approach and a Transformer-based approach with positional embeddings. Existing re-search in this field primarily focuses on models that fully colorize black