Anton Gudiswitz

December 10, 2020

An Algorithmic Approach to Assessing Creativity in Children's Books

For this project I wanted to see if there exists a quantitative approach to assessing visual creativity. Within this paper I will describe the computational method I implemented which was originally published by Elgammal & Saleh. I will then discuss my experiment assessing creativity of the covers of children's books to question how creativity correlates to popular appeal.

1. Computational Method

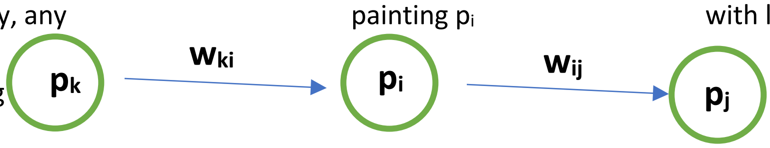
My approach is a recreation of the method used by Elgammal and Saleh [1]. Within their 2015 paper Elgammal and Saleh describe a general formula for assessing creativity for art of any kind, and then apply it to an archive of 62,000 significant western paintings. This approach defines the creativeness of a particular work as a combination of its originality and how influential that work has been. With this definition, the creativity of a specific work of art can be assessed by quantifying how dissimilar it is from the works that preceded it and how similar it is to subsequent works. This framework described could actually be applied to any creative domain, assuming there exists some function that can score the similarity between two works. Here I will be describing it with the example of paintings.

Given a set of paintings P = $\{p_1, p_2, \ldots p_n\}$ and the dates of their creation, our goal is to assign a creativity score $C(p_i)$ to each painting p_i.

First, we create a directed network that reflects the influential value of paintings in the set. Let each painting p_i have a set of n weights W_{ij} that correspond to every other painting in the set. Each weight w_{ij} reflects the influence of painting p_i to the corresponding painting p_j:

a) If p_i was created after p_j, then weight w_{ij} is zero, because art cannot influence art created before it.
b) If p_i was created before p_j, then weight w_{ij} is equal to $S(p_i, p_j)$, which is a function that takes two paintings and produces a positive score of similarity, with higher values indicating higher similarity.

Figure: Illustration of three paintings in the network and the weights connecting them



Currently, any painting p_i with low incoming weights indicates that p_i is novel, and high outgoing weights imply that p_i is influential. We now need to apply several steps to convert the weights in this network into a single score for each painting.

Limit the weights: Because every painting has a weight connected to every other painting, this will result in a biased network with the oldest paintings having a ton of outgoing connections and the newest paintings having a ton of incoming connections. To resolve this for each painting p_i we only keep the K highest weights (where K is some arbitrary constant) and set the rest to zero. Now the remaining weights correspond to the K paintings created after p_i which are the most similar to it, and therefore most likely to have actually been influenced by it.

Balancing Function: This will define what are considered high and low weights. For each weight w in our network if w > 0 then subtract some constant *m*. Otherwise if w =0, then do not change it.

Reverse Negative Weights: The balancing function introduced negative weights. This will cause errors in our next step. To solve this, we reverse all the negative weights (e.g., if w_{ij}=-5 and w_{ji}=0, now w_{ij}=0 and w_{ji}=5,). This keeps the relationship that low incoming weights and high outgoing weights denote high creativity, with the intuition that a negative outgoing weight is equivalent to a positive incoming weight.

Compute the scores: For each painting, the creativity score can be determined with the following formula:

$$C(p_i) = \frac{(1-\alpha)}{N} + \alpha \sum_j \tilde{w}_{ij} \frac{C(p_j)}{N(p_j)},$$

The creativity of a painting p_i is dependent on the sum of all outgoing weights times the creativity of score of the outgoing painting.

The constant α reflects the chance that similarity between paintings is not due to due to influence by prior works. Paintings could also be similar by chance, or because they are part of the same art movement.

N and *N(*p_i) are normalization terms. This indicates that the influence of a painting is split between all outgoing paintings based on weights.

My actual calculations were done with the derived non-recursive form which is the following:

$$C^* = \frac{(1-\alpha)}{N}(I - \alpha\tilde{\tilde{W}})^{-1}\mathbf{1}.$$

I frankly do not understand how Elgammal and Saleh derived this closed form but given a stochastic matrix of weights this formula outputs a vector of creativity scores of each painting.

2. Children's Book Experiment

For my experiment I wanted to apply this method to the covers of children's books. I chose children's books specifically because they are often known for their rich illustrations, fantastical whimsy, and creativity. They also tend to be simpler compositions with less detail than a typical piece of classical Western art.

For my data I collected the covers from Time Magazine's 2016 list *100 Best Children's Books of All Time* [2]. I chose this list because it covered a broad range of publication dates (1902 – 2013), contained many titles I recognize, and the books are critically acclaimed, making them more likely to have influenced each other. Each cover is taken from the original edition.

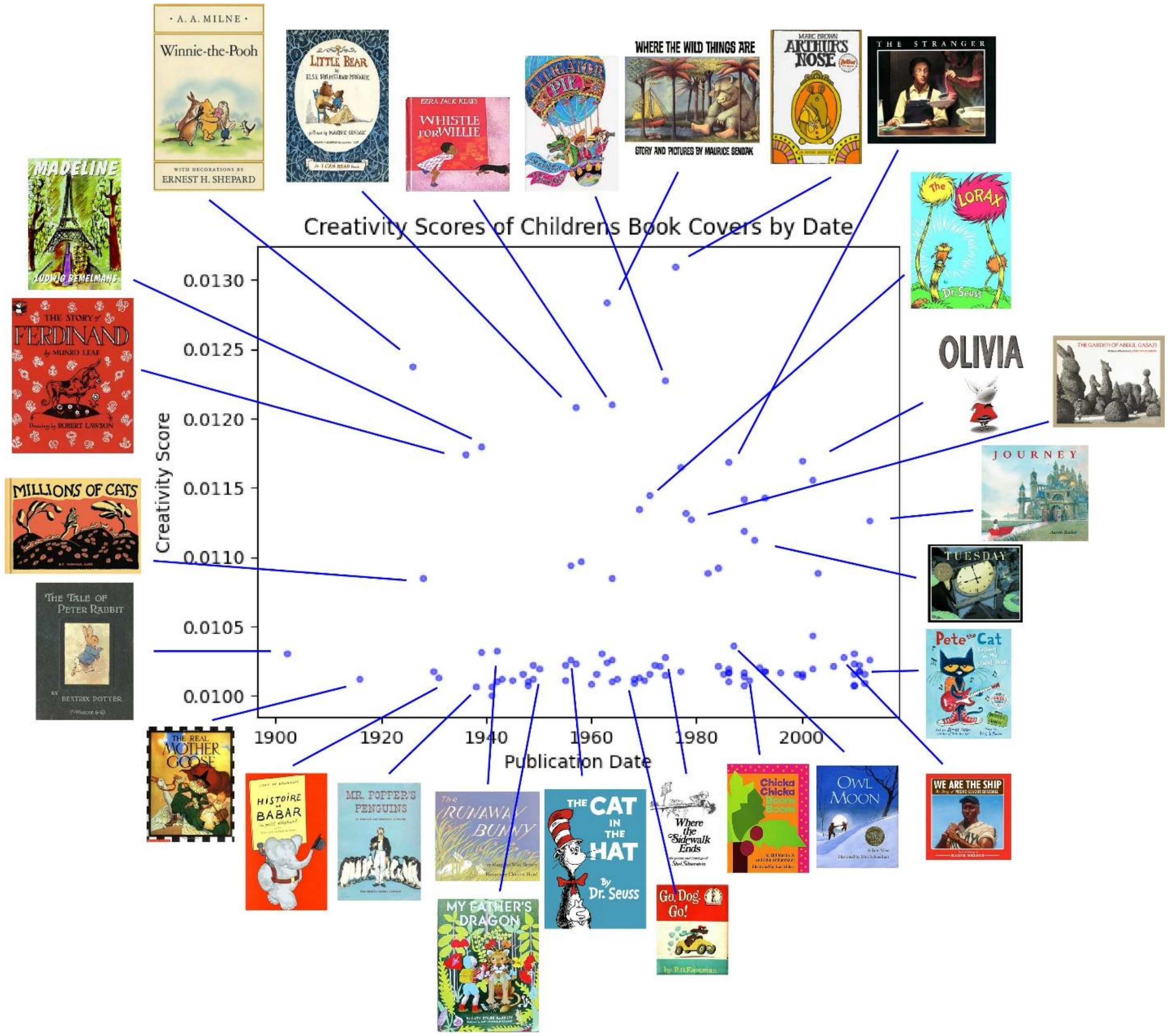
I pursued three research questions with this experiment:

1) Does this method provide a valid assessment of creativity?
2) Do creativity scores correlate with current popular appeal?
3) What books, authors, or illustrators have the most creative covers?

To answer question one, assessing the validity of the algorithm, I would look at outliers in creativity myself and judge based on my own experience with children's books whether I agree with the algorithm. While this is hardly a rigorous method, I expected that the implementation of the similarity function would be quite subjective, so as long as there was not any heinously incorrect seeming scoring then I would entertain the findings.

To test whether creativity scores correlate with popular appeal, I would compare the creativity score of each book with its average rating on Goodreads. With most books in my dataset having over 100,000 ratings this seems to be a reasonable metric for popular appeal. I would expect creativity score to have some correlation with popular appeal, because children's book illustrators would more likely take influence from books they like.

The last question is just a matter of examining the results of the algorithm. I would expect an author's creativity scores to decline with subsequent works, as they would likely keep the same style and decline in originality.



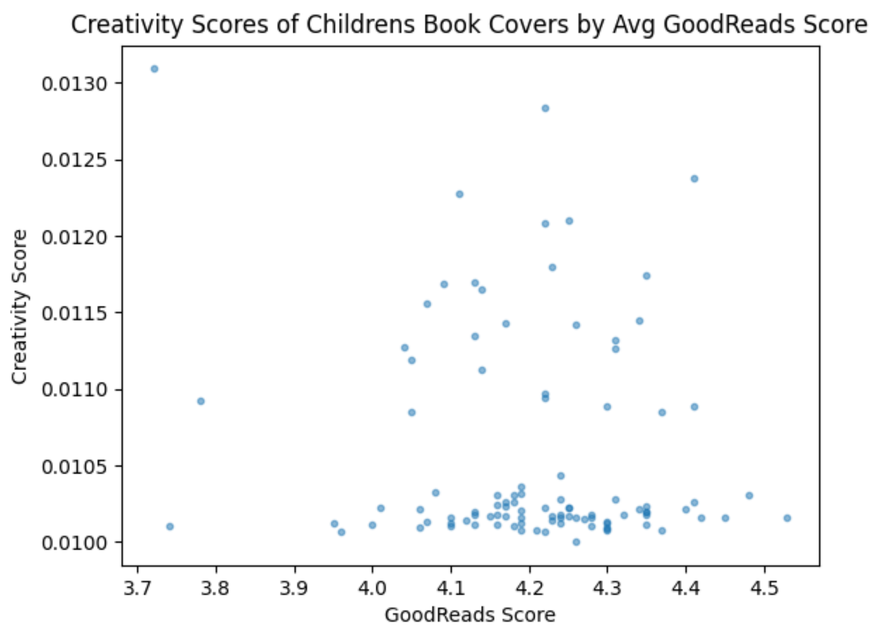
Results

Validity**:** According to this algorithm, the cover art for *Arthur's Nose, Where the Wild Things Are, Winnie-the-Pooh,* and *Alligator Pie* scored highest, while the majority of books fall into a cluster of low scores, among them classics such as *The History of Babar, The Cat in the Hat,* and *Go, Dog Go!*. Just from looking at the data, it is fairly difficult to notice visual trends separating high scoring books from low scoring ones. They might have more varied textures compared to the low-ranking books, which tend to be flat. I would say that all the high scoring books appear fairly original, although the same can be said for many low-and mid-scoring books as well. A high score for *Winnie-the-Pooh* seems to make sense because despite its plain mostly monochromatic look, it is the first of its kind, so all the plain looking covers that follow it indicate its influence. However, the highest scorer *Arthur's Nose* with its centered title and illustration is not all that different from *The Tale of Peter Rabbit, Winnie-the-Pooh, and Little Bear*, at least not enough to justify the #1 spot.

There are many limitations that likely affect the validity of this specific experiment.

- The data was limited to only 100 books; more entries would more rigorously assess influence and originality.
- The images used as data were of inconsistent resolution and dimensions. I tried to find images in the range of 500p X 500p - 300p X 300p, but many do not fit these bounds, and a few are photographs of covers with slightly altered colors and resolution.

- The similarity function I used to compare images is likely flawed in this context. In the original paper, images are compared with a specific computer vision software that uses Classme features, which has been shown to outperform other methods for style classification [3]. I was not able to get a Classme software to compile (despite 10 hours of attempts), so instead I used DeepAi.org's image similarity API [4]. Unfortunately, this algorithm is more suited to assessing visual similarity in photographs, so it is less effective towards inspecting illustrations. Important characteristics such as font, scene context, metaphor and humor likely went entirely unnoticed.
- Artistic influences are most certainly not limited to previous children's books. Many children's books take influence from art styles outside of children's media. This is reflected in the formula by α, but this parameter is set more-or-less by guessing without a clear way to verify.



Popular Appeal: As can be seen in the chart above, there are not any strong correlation of our creativity score values to book ratings on GoodReads. Interestingly the book with the allegedly most creative cover (*Arthur's Nose)* had the lowest GoodReads score. Some reasons might not have worked:

- Our creativity score assessments are invalid
- We are judging books solely by their covers. The story and contents of children's books, not to mention nostalgic value is likely what really makes us like them.
- Cover art often changes in subsequent publications, so the influence can be lost.
- Goodreads average ratings were in a relatively small range. With more books included in the data this might be remedied, but also more obscure books have less reliable ratings.

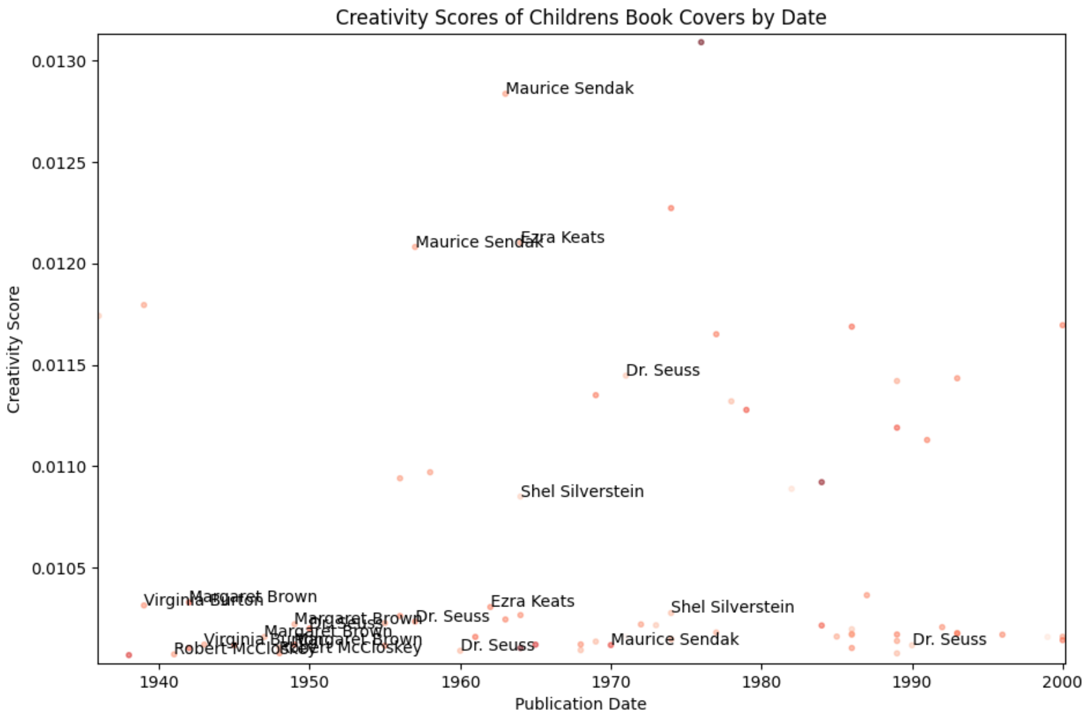


Figure above: Creativity scores listed with the book's author. Color denotes average GoodReads Score (darker = higher rating)

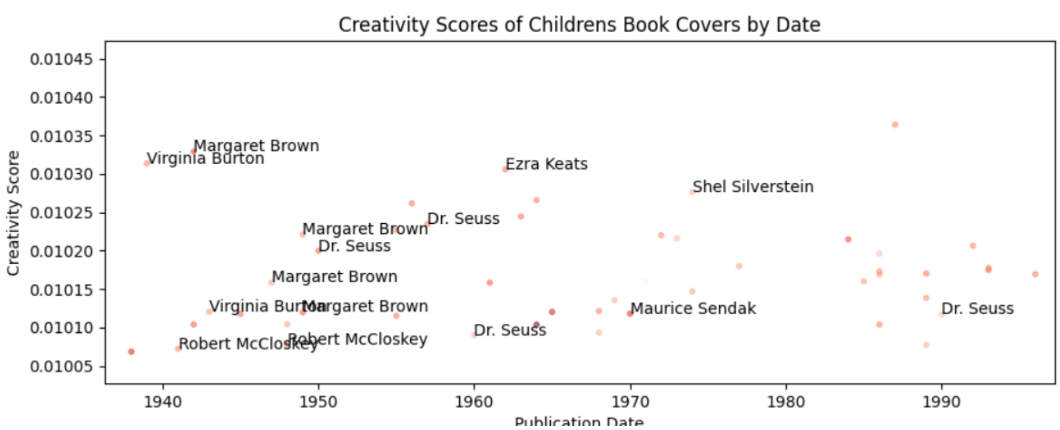


Figure above: A closeup of low scoring book covers and their authors from the previous graph.

Authors/Illustrators: Lastly, we can peak at some of the more well-known children's book authors. Dr. Seuss was the most represented author in the data with 5 appearances. He was followed by Margaret Wise Brown at 4, Maurice Sendak at 3, and Shell Silverstein, Robert McClosky, Virginia Burton, and Ezra

Jack Keats all at 2. Authors do appear to be somewhat clustered, but it is likely due to general trends in the data and not because of authors. Maurice Sendak stands out for having two highly creative works (*Where the Wild Things Are* and *Little Bear)*. As expected, the most creative work of most authors is their first, with subsequent works having lower scores (with the notable exception of Ezra Jack Keats). This offers some support to the validity of our algorithm.

3. Conclusion

To improve my experiment with children's books, I would suggest incorporating more cover-art specific features into the similarity function. Specifically looking at font, textures, and 2-d composition would likely improve the validity of the similarity function considerably. Additional features of the book such as plot or characters could be incorporated to look beyond just cover art.

Personal assessments of a work's creativity are greatly dependent on our subjective experiences, at least in art. Familiarity with the artform, knowledge of its history, and the specific artistic criteria considered would all certainly affect an assessment. While this algorithmic method is far from objective, it is capable of identifying significant creative works without any knowledge of history or art. This could be applied in art history to identify potentially overlooked significant creative works by comparing artworks with specific criteria.

References:

[1] Ahmed Elgammal and Babak Saleh. Quantifying creativity in art networks. In *Proceedings of the 6th International Conference on Computational Creativity*, 2015.

[2] 100 Best Childresn's Books. 2016. https://time.com/100-best-childrens-books/

[3] Lorenzo Torresani, Martin Szummer, and Andrew Fitzgibbon. Efficient object category recognition using classemes. In *ECCV*, 2010.

[4] DeepAi Image Similarity API. https://deepai.org/machine-learning-model/image-similarity

Code and Data: https://github.gatech.edu/Agudiswitz3/Creativity-Assessor.git