Evaluating image matching methods for *Magic: The Gathering* playing card identification

Ada Weinert Ravn
Faculty of Mathematics and Natural Sciences, University of Oslo

(Dated: May 29, 2022)

In this project three popular feature descriptor algorithms, namely Scale Invariant Feature Transform (SIFT) and Oriented Fast and Rotated Brief (ORB) are used for experimental object detection of *Magic: The Gathering* playing cards. The cards are considered at various angles, positions and degrees of clarity and also accounting for additional factors such as glare, foiling, printing and alterations. The results are presented in terms of a comparative analysis or the recognition accuracy considering both the card name exclusively as well as name and printing together.

I. INTRODUCTION

The over 20000 unique *Magic: The Gathering* playing cards add up to over 60000 cards including different printings of cards, and that just including English versions. The ability to distinct those cards and printings is important both to the players and collectors alike, with the value of the card changing from edition to edition. Massive collections need to be catalogued digitally, and name of the card might need to be found if not seen in English. Additionally looking up a card in the store to compare with online stores availability and prices is greatly beneficial to many players. Most of the current detection algorithms are only available for the online versions of the game, or for a very limited subset of cards.

To solve the issue we will attempt to apply the problem to the popular feature descriptor algorithms: Scale Invariant Feature Transform (SIFT) and Oriented Fast and Rotated Brief (ORB). The algorithms have been previously used for projects like identifying the book covers[1] or applications for distorted images[2]. This indicates the possible application here as well.

We will first briefly explain the methods used in this project, covering the card characteristics, feature detection algorithms, feature matching algorithms and brief evaluation metrics. This will be followed by the description of the source data, implementation and finally the presentation of the results. We will conclude with a brief discussion and summary the findings.

Figure 1. Tested cards: *Wrath of God*, *Damnation*, *Lightning Bolt* and *Thoughtseize*. [3]

II. METHODOLOGY

A. Card characteristics

We have selected a focus on 4 cards, as seen in Figure 1, to perform the test of our algorithm. Different cards have different artwork and text, but sometimes cards might appear similar to the naked eye.

The differences in printing can vary from an entirely different image, frame and text box, to minor difference in shade and the set symbol (symbol on the right in the middle of the card). This is represented on Figure 2.

Figure 2. Example of differences between printings. [3]

B. Feature detection

1. *Scale Invariant Feature Transfor (SIFT)*

"SIFT feature detector, proposed by Lowe solves the image rotation, affine transformations, intensity, and viewpoint change in matching features." [2] "The SIFT algorithm consists of four steps. The first step is to search potential interest points by detecting extrema in the difference of Gaussian pyramid (DoG). In the next step, key points are localized and filtered by discarding lowcontrast and edge-response candidate points. Each remaining key point is then associated with an intrinsic orientation based on local image gradient directions. Lastly, local image descriptor is generated for each key point based on image gradient magnitude and orientation." [1]

2. *Oriented FAST and Rotated BRIEF (ORB)*

"ORB is a fusion of the FAST key point detector and BRIEF descriptor with some modifications." [2] "The FAST algorithm (features from accelerated segment test) is a corner detector. It uses a circle of 16 pixels to classify whether a point is a corner. ORB extracts FAST corners from a multi scale image pyramid to produce scaleinvariant features. It also applies Harris measure to find top N points. ORB uses a technique called intensity centroid to assign orientation to each key point. Moments are computed to improve rotation invariance. The BRIEF descriptor (Binary Robust Independent Elementary Features) is a bit string description of an image patch constructed from a set of binary intensity tests. To achieve invariance to in-plane rotation, ORB computes a rotation matrix using the orientation of patch then the BRIEF descriptor are steered according to the orientation." [1]

C. Feature matching

The feature sets are matched using the simplest method of taking the descriptor of one feature in first set and matching it with all other features in second set using some distance calculation. Usually the Euclidean distance is applied for the SIFT and SURF features we use the Euclidean distance, and the Hamming distance for the ORB features. This is known as the Brute Force Matching.

It is further possible to return k best matches instead using the k-Nearest Neighbours algorithm and then select the better match using the Lowe's ratio test. It is a method that filters feature matches by eliminating matches when the second-best match is almost as good. This helps remove many false matches - decreasing the noise.

D. Evaluation metrics

We inspect the accuracy both as general card recognition, useful for identification of card for the purpose of importing the ruleset or legality, as well as detailed card and printing recognition which can be used for automatic identification of cards in collection.

III. SOURCE DATA

A selection of photos of the four following cards have been gathered as seen in Figure 1: *Thoughtseize*, *Damnation*, *Wrath of God* and *Lightning Bolt*. The examples include tests related to photo quality, card recognition difficulty and a combination of both. In the first group

we include factors such as rotation, obstruction, glare, different lightning conditions and cropping. In the second group we include complications such as foiling (usually rainbow reflective film added to the card), different language versions, added signatures, alters (decorative painting over parts of the card) and printings. All card names and printings are identifiable by human eye.

The set of photos is matched against a test set of all available printings for each of the four cards as well as a control group of similar looking and random cards for a total of 100 match-able cards. The cards images are imported from Scryfall[1] in the English version using the 'large' format. The photographs were gathered from various sites allowing for reselling cards[2][3]. Most have since been taken down (sold).

IV. IMPLEMENTATION

The models were implemented and evaluated using Python's OpenCV Library [4]. The default initialization parameters are used for each model.

V. RESULTS

I would like to prefix saying that both algorithms performed very well on clean images and with small amount of possible target images. As a result the testing was continued into a combination of challenging images with a somewhat larger target group instead. It was not expected that all of the cards chosen would be possible to accurately classify, so the main goal is to try to identify potential problem areas and overall better algorithm - if any.

Method	Correct	Partially correct	Misclassified
ORB	30/50	8/50	12/50
SIFT	41/50	5/50	4/50

Card matching using features from the whole card. Partially correct indicates correct card type but incorrect printing.

First a test against the whole card image was performed on the test data. A few patterns can be observed from the get-go: the inclusion or crop of the background and rotation of the cards in the photos do not seem to affect algorithms' ability to match the images for either algorithm. We further note that the two algorithms appear to struggle with a different subset of images.

Figure 3. Examples of misclassified cards using the SIFT algorithm.

Although the SIFT algorithm is overall pretty accurate it does seem to particularly struggle with identifying the card *Wrath of God*, characterized by low contrast in the image frame. It has been consistently mistaken for cards with a similar characteristic, here in the case of 3 out of 4 mismatched cards as seen in Figure 3. The last misclassified card is a moderately altered copy of *Damnation*, which proved to be problematic for all of the tested versions of the program. Overall 46 out of 50 cards have been at least partially correct, i.e. have had correct card identified, looking apart from the printing, which is a very encouraging score.

Next we look at the partially misclassified cards as seen in Figure 4. We seen an issue with glare and further issue with the above mentioned *Wrath of God*, overall only half of the tested pictures of this cards have been correctly identified by the algorithm. Minor issue with unfortunate foiling reflection is also noticed.

The ORB algorithm, although much faster, is not nearly as accurate. It struggles with dark signatures, i.e. a black signature over the text box, and also - to a degree - card alterations which have relatively high contrast (relative to the same area in an unaltered image). We note that small alterations and silver signatures do not have a notable effect on the accuracy of the prediction.

[1]Link to documentation: https://scryfall.com/docs/api
[2]https://www.cardmarket.com/en/Magic
[3]https://www.ebay.com/
[4]Link to documentation: https://pypi.org/project/opencv-python/

Figure 4. Examples of partially correct cards using the SIFT algorithm.

Figure 5. Examples of misclassified cards using the ORB algorithm.

A large subset of cards including foiling has also been entirely misclassified, though we note that foiled cards with unique (promo) frames have not been affected by this. Additionally this algorithm does not handle photos where less than half of the card is visible in the frame, misclassifying all of the included examples. 30 out of 50 photos of the cards have been correctly classified, with examples of erroneous classifications seen on Figure 5.

The cards that have been classified with an incorrect edition follow a similar pattern of alterations and foiling causing issues as seen in Figure 6. Additionally cards partially covered by a shadow have had wrong editions assigned to all of them. We note an issue with a single card *Wrath of God* with apparently clean picture not being classified correctly. Further testing is required to establish why that image in particular was misclassified.

Method	Correct	Partially correct	Misclassified
ORB	6/50	10/50	34/50
SIFT	17/50	28/50	5/50

Card matching using features from the artwork only.

Afterwards a test using only the artwork in the card for the feature comparison. Naturally only results for the partially correct matches were expected as many printings share the same artwork, which, at most, has contrast and color balance altered. The ORB algorithm didn't perform well here at all, but we note that the SIFT algorithm performed almost as well as for the previous test with 45 out of 50 cards being correctly partially classified.

We can compare the partial results to the previous SIFT test. We see that two cards previously misclassified remain misclassifed in this approach. The other two - the two versions of *Wrath of God* - have now been correctly classified. Two of the cards that were partially classified are not completely misclassified - a poor quality image in a frame, and a tilted image with a glare. Finally one image with minor alteration to the frame (outside of the match-able area) is now also misclassified for unknown reasons.

VI. DISCUSSION

Although slower, we notice that the SIFT algorithm performs very well in terms of accuracy - even considering only the initial card identification. It is very accurate,

Figure 6. Examples of partially correct cards using the ORB algorithm.

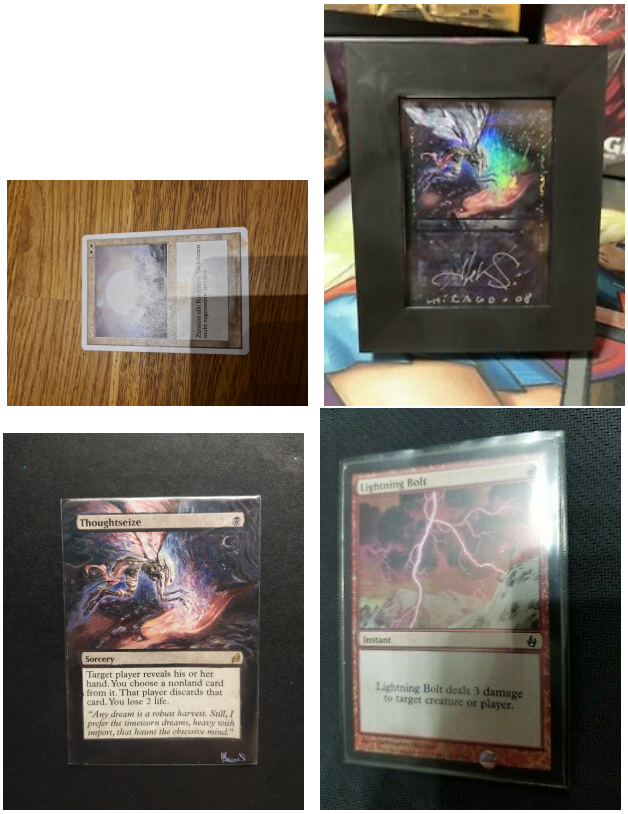

Figure 7. Examples of misclassified cards using the SIFT algorithm - artwork only.

even for very unclear photographs and handles most of the glare, foiling and even alterations. We notice the issues with cards with low contrast in particular, even for very clear photographs - which is something to take under consideration. This could be solved by focusing only on the artwork as then the algorithm performs much better for low contrast images.

The ORB algorithm is very fast, fast enough to potentially use for card recognition for much larger subsets of cards than 100. What it struggles with is prediction accuracy. We see big issue with foiled and altered cards, as well as partially hidden images.

If we can find an easy and fast way to identify whether the card is altered and/or foiled - either by user identification or an additional detection algorithm the fast ORB might still be usable for general identification of cards. After narrowing the card down to a correct estimation we can then run the slower SIFT algorithm to compare it to all available printings for that specific card - narrowing it down to on average 10 cards for the commonly played cards.

It would be interesting to also try the SURF algorithm, which is the faster version of SIFT, to see if it would be a good competitor to this solution. Unfortunately it ended up being beyond the scope of this project.

Overall the results were satisfactory for what was being tested, but another solution to the problem of identifying cards from a large selection might be necessary,

VII. CONCLUSION

In this project we have studied the relative prediction accuracy of image matching models ORB and SIFT when applied to *Magic: The Gathering* playing cards. We have considered different printings, editions, foiling, glare from the photo, rotation, picture at an angle, alterations and signatures. The photographs were then matched to a database of card images for the purpose of identification.

We have discovered that the algorithms have different strengths and weaknesses: ORB is fast but doesn't handle foiling and alterations, SIFT is slow but very accurate. A combination of approaches might be necessary to solve the issue of efficient card identification.

Further work should include testing different algorithms as well as a combination of different solutions.

REFERENCES

[1] B. W. Rabie Hachemi, Ikram Achar and M. S. A. Mebarek, "Evaluating image matching methods for book cover identification," *Evaluating image matching methods for book cover identification*, 2001.

[2] S. P. Ebrahim Karami and M. Shehata, "Image matching using sift, surf, brief and orb: Performance comparison for distorted images," *Image Matching Using SIFT, SURF, BRIEF and ORB: Performance Comparison for Distorted Images*.

[3] "https://scryfall.com/docs/api/images," 2022.