Computational Linguistics in the Netherlands Journal 12 (2022) 253-267 Submitted 12/2022; Published 12/2022

Automatically Interpreting Dutch Tombstone Inscriptions

Johan Bos* JOHAN.BOS@RUG.NL
Cristian A. Marocico** C.A.MAROCICO@RUG.NL
A. Emin Tatar** A.E.TATAR@RUG.NL
Yasmin Mzayek** MZAYEKY@GMAIL.COM

* *Center for Language and Cognition, University of Groningen*

** *Center for Information Technology, University of Groningen*

Abstract

Digital preservation of tombstones is important in the context of cultural heritage but a costly process. We propose a way of automatically reading tombstone inscriptions with the aim of assisting human annotators and data curators. Our method comprises a pipeline of dedicated components where the input is an image of a tombstone, and the output is an interpretation (represented as a directed acyclic graph) comprising the names of the deceased, dates of birth and death, places of birth and death, and biblical references. The three main components in the pipeline are (1) Label Detection, (2) Optical Character Recognition (OCR) and (3) Semantic Interpretation. The Label Detection component uses an off-the-shelf deep Learning algorithm trained on tombstone images to detect the bounded boxes and labels for the entities mentioned above (names, locations, and dates). The OCR component then takes in each of the detected labels and recognizes the text contained therein. Finally, the interpretation component performs post-processing, normalizes dates and places, and puts all the information together into a meaning representation coded as a directed acyclic graph. There are several challenges that need to be addressed, such as correcting OCR errors, interpreting unusual segmentation of words, recognising abbreviations, dealing with multiple languages, and accommodating different notational variants of dates. The system, the first of its kind, is developed and evaluated with the help of an annotated corpus of 1,100 tombstone inscriptions. Evaluation is carried out by calculating graph overlap of the system output compared to gold standard. The results are encouraging, with an F-score of 67% outperforming a random baseline of 40%.

1. Introduction

The digital conservation of tombstones has seen an increase of interest in the context of cultural heritage preservation and sparks interest of many disciplines including anthropology, archaeology, history, linguistics, philology, sociology, and theology (Saller and Shaw 1984, Eckert 1993, Veit and Nonestied 2008, Long 2016, Bos 2022). This process has also become more urgent as tombstones or graveyards disappear because of weathering (Figure 1) and urban or political developments (Matthias 1967, Streiter and Goudin 2011).

The digitalisation process of funerary material is costly and time-consuming, as it involves taking high-quality pictures of the gravestones, performing optical character recognition (OCR) on the images, and interpreting and correcting the OCR results. It is the latter part of this process that is the topic of this article.

Our aim is to assist human annotators and data curators by proposing a way of automatically reading and interpreting tombstone inscriptions. An automatic tombstone parser has the potential to increase the efficiency of human annotation, where the idea is that the system produces a draft representation that can be verified and corrected by a human annotator. The system that we aim to develop takes as input a photograph of a tombstone, and outputs a meaning representation with

Figure 1: Weathering threatens the preservation of tombstone inscriptions. Zuiderbegraafplaats Groningen, 2 October 2019 (image by Johan Bos).

the entities that appear in the tombstone inscriptions (persons, locations, and dates) and relations between them (dates and places of birth and death). Figure 2 illustrates this idea.

Given a dataset with more than a thousand tombstone images annotated with meaning representations, our aim is to investigate how far off-the-shelf components (dedicated to object identification and optical character recognition) will help us to build a tombstone parser with adequate performance. First, in Section 2 we present related work. In Section 3, we give an overview of challenges that arise when parsing tombstone images. Then, in Section 5, we explain what the major components of our tombstone parser consists of and how they were trained on the data. Finally, in Section 6 we present and discuss our results.

2. Related Work

Our work has points of contact with the XML-based annotation scheme for graveyards and tombstones proposed by Streiter et al. (2007). Whereas this scheme focuses on the physical properties of tombstones, we aim to analyse the content of tombstone inscriptions. This content includes people, dates, locations, symbols, religious references that are mentioned on tombstones and the relations between them (temporal relations, family relations, occupational relations).

From a technical point of view, our work also bears strong similarities with the research by Franken and van Gemert (2013), who developed a system for automatically recognising and interpreting ancient Egyptian hieroglyphs from photographs, based on a dataset of about 4,000 hieroglyphs.

But as far as we are aware, we are the first to develop a semantic parser for tombstone images, i.e., a system that takes as input a photograph of a tombstone, and outputs a meaning representation. Semantic parsing is one of the key tasks in computational semantics (Blackburn and Bos 2003). The meaning representations that we target are similar in format to AMR, Abstract Meaning Representation (Banarescu et al. 2013). Several high-performance semantic parsers have been developed for

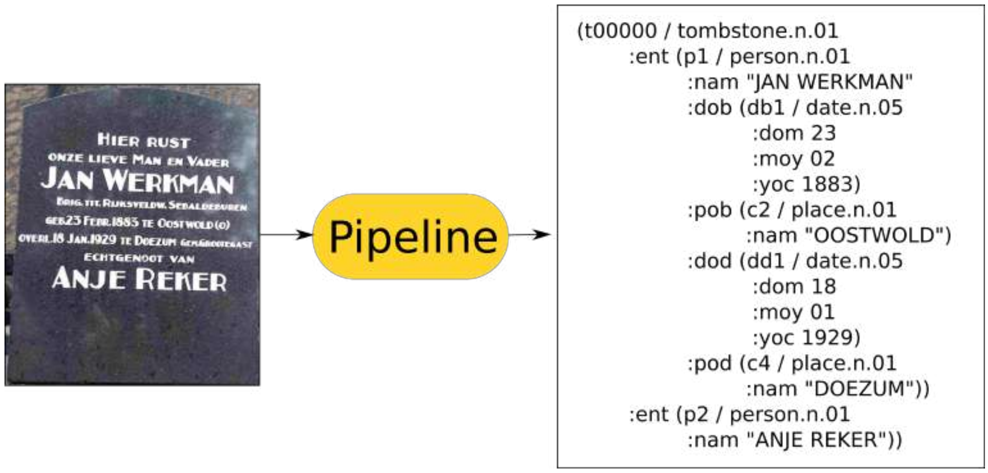


Figure 2: The idea of automatically interpreting a tombstone image.

AMR (Artzi et al. 2015, Van Noord and Bos 2017, Lyu and Titov 2018). However, these parsers take written sentences as input, not images of tombstones. Parsing text on photographs offers many new challenges for semantic parsing.

3. Challenges

Semantic parsing applied to images of tombstone inscriptions raises a number of new challenges that are not present in the traditional way semantic parsing is applied to digitally clean text, such as newspaper stories or scientific articles, that are often already available in electronic format (Figure 3). These challenges can be grouped into the following categories:

- image: properties of the image, such as limited contrast between stone and inscription, variation in background and foreground colour, and a mixture between symbols and text;

- presentation: the way information is presented on the stone, e.g., the variety in fonts, size, spacing, word and line breaks, and use of columns;

- linguistic: use of pronouns, unclear boundaries of names and locations, many different notations for dates, coordinate structures using parenthesis, long-distance dependencies.

The task of optical character recognition (OCR) needs to deal with the surface variation that can be found on stones to detect regions on the stone that are relevant for semantic interpretation. The way information is presented can be problematic for segmenting the text into entities (names of people, places, dates, and biblical references). The output of OCR is rarely perfect, so a further challenge exists in detecting and correcting OCR errors of names, places, and dates.

The use of pronouns and multiple occurrences of the same person on a stone is challenging for entity linking and semantic role labelling. Connections between entities are sometimes difficult to establish because they are not always local. For instance, an inscription could read: "In memory of Mary Brown, widow of Stephen Black, born April 1st, 1925" could be erroneously interpreted to mean that Stephen was born in 1925, rather than Mary, on whose tombstone this inscription appears. Thus, proximity in the text between two entities does not necessarily imply a connection between them (Figure 3).

Figure 3: Examples of tombstone images that pose challenges for automatic interpretation: reduced spacing between words, long-distance dependencies between names and dates, patronyms, locative pronouns, dates on separate lines, poor contrast and mixture of fonts.

4. Dataset

We use the dataset collected and created by Bos (2022), which consists of about 1,100 high-quality tombstone images collected from forty public cemeteries in the three northern provinces of the Netherlands (Friesland, Groningen, and Drenthe). The images are cropped towards the outside borders of the stone, so they differ in size and height-width ratio. Each tombstone image is paired with a meaning representation that takes the form of a rooted directed acyclic graph. The graphs are formatted using the PENMAN notation (Kasper 1989, Bateman and Paris 1989, Bateman et al. 1989, Bateman 1990), a convenient format that is used in Natural Language Processing and for which evaluation software has been developed (Cai and Knight 2013).

The semantic graphs are elaborate and contain considerable detail in describing concepts and relations between them. The nodes of the graphs describe entities as WordNet (Fellbaum 1998) synsets (e.g., male.n.02, spouse.n.01, frond.n.01, and so on) or constants (names of people, grounded locations, date expressions). The edges of the graph are semantic relations, using an inventory of over twenty different relationships.

For the purposes of our research we simplified the graphs, leaving out non-essential concepts and relations and collapsing several concepts under a common unifying concept. This simplification can be summarised as follows:

- male and female persons (male.n.02 and female.n.02) are mapped to person.n.01;
- different types of locations (city.n.02, village.n.02) are mapped to place.n.01;
- dates expressions are decomposed from a single string format `YYYY-MM-DD` to their three components: `YYYY`, `MM`, and `DD`;
- nodes representing symbolism, professions, family relations and biblical references are removed from the graph.

This simplification is motivated by the divide-and-conquer principle. The problem at hand is complex of nature but can naturally be divided in smaller sub-problems. By mapping male and female to a super-class we set the problem of gender identification aside. The same holds for the different classes of locations. The dates are broken down in order to evaluate accuracy on a finer level.[1] The recognition of professions, family relations and integration of biblical references[2] is left for future research. The motivation to leave out the recognition of symbols (hourglasses, animals, vehicles, and many more) requires more labelled data and a dedicated approach that go beyond the scope of this research. This leave us with the concepts that are listed in Table 1.

Table 1: Concepts (WordNet synsets) and their meaning in the derived dataset.

Concept	**Meaning**
tombstone.n.01	a stone that is used to mark a grave
person.n.01	a human being
date.n.05	the particular day, month, or year that an event occurred
place.n.01	a point located with respect to surface features of some region

Even after leaving out this amount of information, automatic tombstone parsing remains highly challenging. Many a stone shows more than just a single name and a pair of dates. Table 2 illustrates this, showing that on average there are two persons mentioned on a tombstone (with a maximum of 7!), up to ten different dates, and six different places.

Table 2: Distribution of information on stones in the corpus, divided by class. Numbers refer to the frequency a class appears on a stone.

Class	**Average**	**SD**	**Range**
person	2.06	0.78	0–7
date	2.87	1.36	0–10
place	1.20	1.47	0–6

Because we use several machine learning approaches in our choice of method, we have randomly split our dataset of 1,100 images into traditional training, validation and test splits. Following a 80%–10%–10% rule, we then obtain dataset sizes of 897 (train), 101 (valid), and 101 (test).

5. Method

Given established methods of semantic parsing in natural language processing (NLP), one could think of a straightforward solution to the problem, namely one that adds an OCR component in front of a standard NLP pipeline. This standard pipeline would then perform tokenisation, named entity recognition, syntactic parsing, and semantic interpretation, as usual. However, this method faces two problems. The first problem is that the OCR does not capture graphical but only textual information about the inscription. For example, names on tombstone inscriptions are usually highly visible, being written in a larger and bolder font than the rest of the inscription. This we consider only a small drawback, since the named entity recognition part of the pipeline can still correct for this omission. The second and larger problem lies in the future enlargement of the pipeline with more *non-textual* entities, such as *symbols* that occur frequently on tombstones, which would be completely missed by the OCR.

Taking these issues into account, we have decided on a different approach, with a pipeline that takes as input an image of the tombstone, and outputs an interpretation (represented as a directed acyclic graph) comprising the name of the deceased (nam), date of birth (dob), date of death (dod),

1. Otherwise for a date to be correctly interpreted, *all* of its components (day, month, year) need to be correct. By breaking down dates, a system would still get some credit if it predicts the correct value for one or two of the three components.
2. Actually, the biblical references are accounted for in the first stage of our semantic parser, label detection, but not supported in the components that follow in the pipeline.

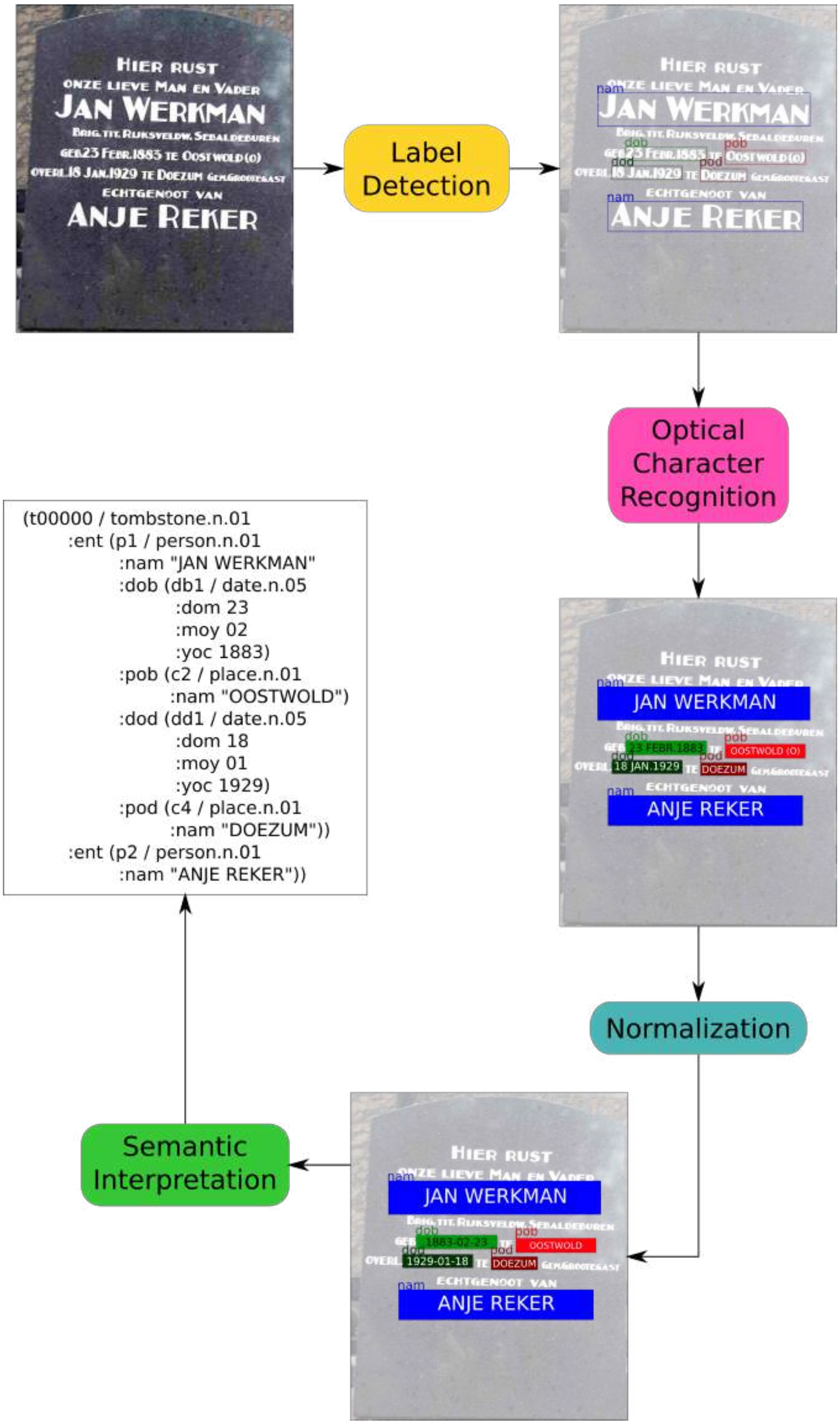


Figure 4: The four components of our automatic tombstone parser with example input and output.

place of birth (pob), and place of death (pod). The pipeline itself comprises the following components (Figure 4):

1. **Label Detection**,
   the task of finding and classifying relevant areas on a stone;
2. **Optical Character Recognition**, or OCR for short,
   the task of transcribing the information of a detected area into text;
3. **Post OCR Processing**,
   the task of verifying locations, correcting names, and normalising date expressions;
4. **Semantic Interpretation**,
   the task of establishing relations between extracted entities.

This is our proposed system for tombstone parsing in a nutshell. In what follows below we describe each of these components in more detail.

5.1 Label Detection

The Label Detection component of the pipeline is implemented by making use of a deep learning network for image classification called YOLOv5 (Jocher et al. 2022). The network is designed to detect multiple objects of different classes in an image. It kills two birds with one stone, as it returns the class of a detected object as well as the coordinates of the rectangular bounding box for that object.

In order to fine-tune the YOLOv5 algorithm on our dataset, we annotated our entire dataset by selecting bounded boxes on the tombstone images and annotating them with six different labels: nam (name of a person), dob (date of birth), pob (place of birth), dod (date of death), pod (place of death), and bib (biblical reference). This was carried out by using the online service RoboFlow.[3]

This platform also allowed us to perform augmentation of the images in the train set and proper scaling (640 × 640 pixels) for input into YOLOv5. Augmentation is needed to enlarge the training set. The augmentation process consists of transforming the images into grayscale, adding noise, blurriness and variable brightness to the images (Figure 5). This way, we have tripled the number of images in the train set to 2,691.

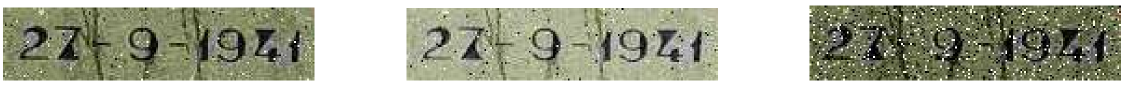


Figure 5: Example of augmented images (cropped) from RoboFlow.

Finally, we have used the annotated train set to train several YOLOv5 models (see the GitHub repo of YOLOv5 from Jocher et al. (2022)), and then tested the model performance on the valid dataset. On a variety of measures ($F1$-score, Precision-Recall curve), the models all performed similarly and reasonably well, so we have decided to use the larger model yolov5x6, with the view that it can be trained in the future on new tombstone images, should they become available.

5.2 Optical Character Recognition

The second component in the pipeline takes care of Optical Character Recognition (OCR): it takes as input the detected area within a tombstone image and outputs the characters that it recognises. Generally, OCR models have two parts; text detection and text recognition models. As described in the previous section, we use YOLOv5 as the text detection tool and feed the detected crops to

3. See https://app.roboflow.com/.

PaddleOCR's[4] customised text recognition model. At this stage we don't use the labels predicted by the Label Detection component.

To customize a character recognition model, we semi-automatically annotated over 15,000 crops returned by YOLOv5 from the images in the training dataset. The annotation has been done semi-automatically by mapping the crops back to PENMAN notation using the detected labels to extract the information on the crop. In case the image contained more than one crop with the same label (i.e., more than one name of a person, or more than one date of birth, and so on), the crop was wrongly labelled by YOLOv5, or the text on the crop was not in the standard form, we corrected or completed the annotation by manual intervention. Special attention was given to date expression, that occur only in normalised form in the gold meaning representation, and not in the way as represented on a stone (Figure 6).

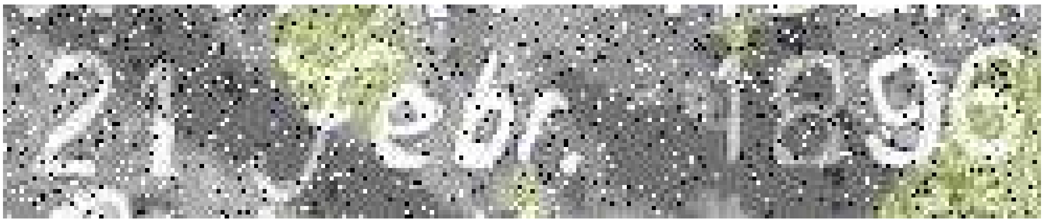


Figure 6: The *date of death* in the above image is recorded as 21-02-1896 in the PENMAN graph. The annotation is manually changed to 21 Febr. 1896 before training a recognition OCR model.

Next, we further trained using the crops from the training images a pre-trained Convolutional Recurrent Neural Network (CRNN) (Shi et al. 2017) which uses MobileNetV3 (Howard et al. 2019) as the embedded feature extractor neural network. The pre-trained CRNN model is trained 72 epochs (about 3,600 iterations) with the crops from the training images. As Figure 7 shows, this number of iterations is sufficient.

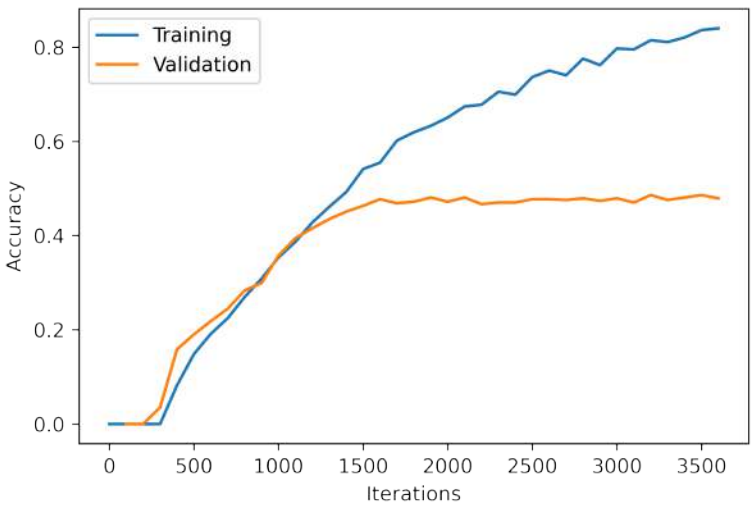


Figure 7: Training performance of the OCR recognition model

4. `https://github.com/PaddlePaddle/PaddleOCR/blob/release/2.6/doc/doc_en/algorithm_rec_crnn_en.md`

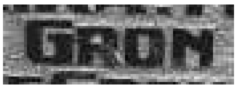
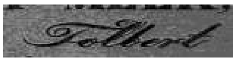
5.3 Post OCR Processing

The *Post OCR Processing* component of the pipeline takes as input the text generated by the *OCR* step, together with the labels generated by the *Label Detection* step. It then performs post-processing on this data, normalising dates, correcting names, and verifying places, in preparation for the final step.

5.3.1 Verifying locations

To verify the places of birth and death, we use the API provided by the GeoNames geographical database.[5] Upon entering the alpha-2 country code (i.e., NL for Netherlands), the names of locations within that country are retrieved from this database to form a list of location names. To deal with the locative pronoun "aldaar" (see Figure 3), we add it to the list as an exception, since it is the most common location placeholder found on tombstones. In the next step the similarity of the OCR-scanned text with a score of less than 0.95 is checked against all the names on the location list. This is done with the help of the Python implementation of the SequenceMatcher algorithm.[6] The location name that is most similar to the one found on the OCR text is chosen as the final result. Table 3 shows a couple of examples.

Table 3: Examples of location crops with the recognised text and their corrections. Corrections are not needed above a threshold of 0.95.

Image	**OCR Result**	**OCR Score**	**Correction**
	GRON	0.9999	n.a.
	TOLLORT	0.902650	TOLBERT
	ALDLAZAR	0.888943	ALDAAR

5.3.2 Correcting names

The main challenges in correcting OCR errors in names are detecting the error and not having a single ground truth as a name might have different spellings. As noted by Nguyen et al. (2021), there are different OCR post processing approaches. In this project, we adopt a lexical approach to address the correction of names. We use the free and open source spell checker GNU Aspell.[7] We include it to our pipeline (Figure 4) using the Python wrapper aspell-python.[8]

The pipeline detects the full names, that is, all given names, the family name, and the affix. In some cases the given names and affix are abbreviated. Our method only processes the cases where the obtained OCR score is below 0.95. It first splits the names into its components and checks whether every component is in the dictionary and if not it returns the word that is closest to it. The advantage of this method is that it only returns meaningful words. On the downside, if the OCR recognition is poor, this method can return irrelevant names which might be less similar than the names predicted by the OCR component (Table 4).

5. See `http://download.geonames.org/`.
6. See `https://github.com/python/cpython/blob/3.10/Lib/difflib.py`.
7. See `http://aspell.net/`.
8. See `https://github.com/WojciechMula/aspell-python`.

Table 4: Examples of person name crops with the recognised text and proposed corrections (here a correct and wrong one).

Image	OCR Result	OCR Score	Correction
	HENK SPGELSTRAN	0.922882	HENK SPOELSTRA
	EROEN BIIL	0.941783	EREN BUIL

5.3.3 Normalising Date Expressions

The dates, no matter how accurately they are recognised by OCR, are normalised into the form `DD-MM-YYYY`, where `D`, `M` and `Y` are all digits. First, a recognised date expression is divided into the day, month, and year components. Days and years are usually representing by numbers, but months are often written in letters. To convert months written in letters into two digits, we use a heuristically built dictionary based on the OCR output of the training data (Table 5). If the attempt to parse the three components of the date fails, then we return the recognised date as delivered by the OCR module.

Table 5: Heuristic conversion rules for turning months into two digits.

Recognised Expressions	MM
JAN, JANUARI, JANUARIJ, JANVIER, I	01
FEB, FEBR, FEBRUARI, FEBRUARIJ, II	02
MAART, MRT, III	03
APR, APRIL, IV	04
MEI, MEIJ, V	05
JUNI, JUNY, JUNIJ, JUN, VI	06
JULI, JULY, JULIJ, JUL, VII	07
AUG, AUGUSTUS, AUGS, VIII	08
SEPT, SEPTEMBER, SEP, IX	09
OCT, OKT, OKTOBER, OCTOBER, OCR, X	10
NOV, NOVEMBER, NOEVEMBER, XI	11
DEC, DECEMBER, XII	12

5.4 Semantic Interpretation

To put all the information together, we arrange the recognised names and places and normalised dates in the order that they appear in the inscription, based on the bounding box coordinates obtained from the *Label Detection* component. We then follow some simple rules to build an acyclic directed graph with root `tombstone.n.01`:

- a detected person P with label `nam` is connected to the root of the graph with the `:ent` relation in the following way: `(X / tombstone.n.01 :ent (Y / person.n.01 :nam "P"))`, where X and Y are new variables.
- a detected place P with label `(pob|pod)` is connected to the person that appears just before it with the following structure: `(:pob|:pod) (X / place.n.01 :nam "P"))`, where X is a new variable.

- a detected date DD-MM-YYYY with label (dob|dod) is connected to the person that appears just before it with the following structure: `(:dob|:dod) (X / date.n.05 :dom DD :moy MM :yoc YYYY))`, where X is a new variable.

These simple set of rules together form a proof of concept, and will not deal with all the challenges mentioned in Section 3. To cover more complex cases the rule sets will need to be extended.

6. Results

We evaluate our tombstone parser in several ways. Because the parser is a modular system, we can evaluate some of the components separately. We first present results on label detection, then have a closer look at the performance of the OCR component, and then wrap it up with by giving the overall results on semantic parsing.

6.1 Label Detection

Figure 8 shows the confusion matrix of the chosen model, `yolov5x6` on the test dataset. It shows high accuracy for all six classes, and there is hardly any confusion between them. The model produces very few *false negatives* for background, meaning that it doesn't ignore many labels that it should classify. On the other hand, it produces a sizeable number of *false positives* for background, wrongly proposing information on a stone that does not belong to the six designated labels.

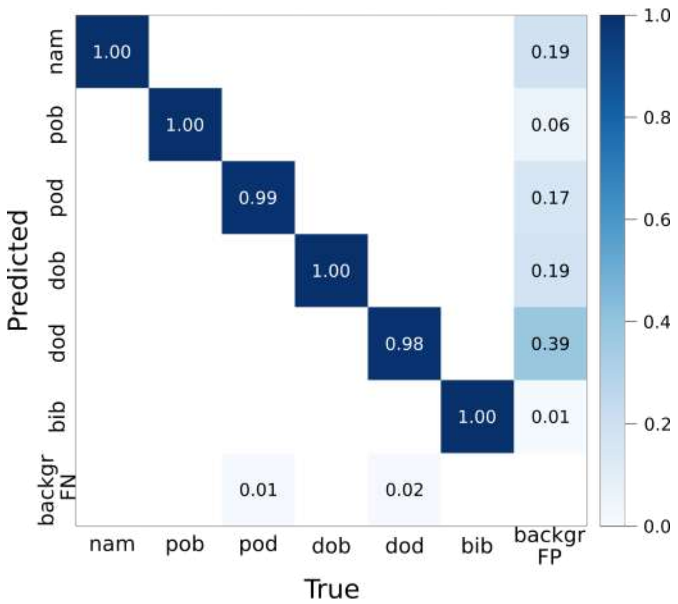


Figure 8: Confusion matrix of the `yolov5x6` model on the `test` dataset.

All in all these results are encouraging. On the one hand it is not surprising that there are not many mistakes in confusing dates from person and location names, because both content and context are likely features that help to do this reliably. On the other hand, it is perhaps slightly surprising that two classes of dates (birth and death) and places (birth and death) can reliably be differentiated from each other.

6.2 OCR performance

The performance of the customised OCR model is tested on 531 crops from the `test` dataset, before and after the post-processing. We calculated the similarity score between the OCR and the exact text on the crops using the SequenceMatcher algorithm which returns a score between 0 and 100.

Before post-processing, the mean similarity score was 69.53. The OCR model predicted 129 of the text or 24.29% fully correct (i.e. 100 similarity score). We can see from Fig. 9, OCR model performed the best on locations and the worst on biblical entities.

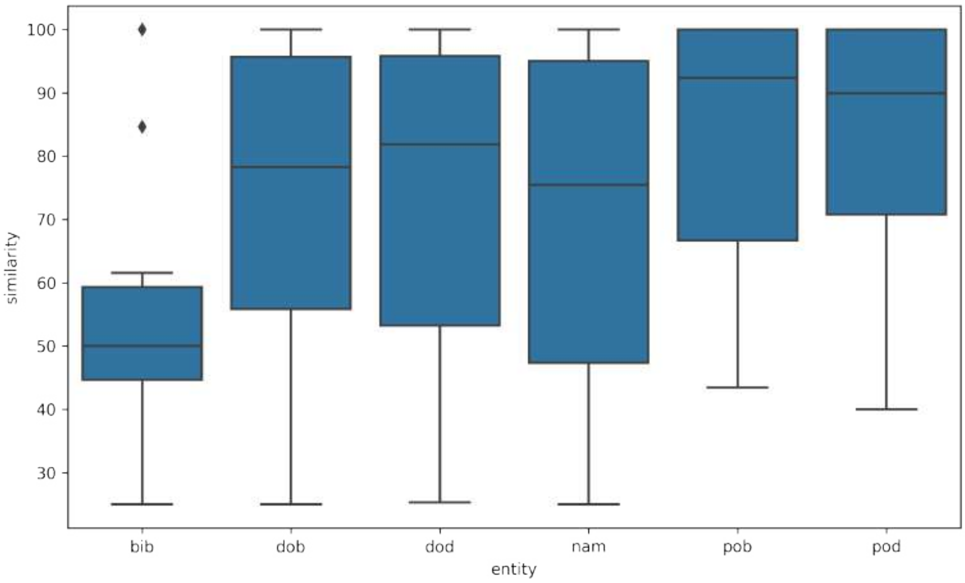


Figure 9: Box plot showing the similarity distribution across different entities.

After post-processing of the names and the locations as explained in 5.3, the mean similarity score dropped to 68.55. However, the number of 100% matched text raised to 153 or 28.81%. This suggests, if OCR text is close to the original text, then post-processing makes it better. Otherwise, it reduces the similarity. If looked at in detail, the corrections of the locations contribute with 23 more 100% matched texts to the increase in the number of total 100% matched texts.

6.3 Interpretation (Final results)

The tombstone parser that we developed was evaluated on a corpus of 1,100 tombstone images. Evaluation was carried out by calculating graph overlap of system output and gold standard. This task was performed by making use of the off-the-shelf tool *Smatch* (Cai and Knight 2013). Smatch is a tool originally developed for evaluating semantic parsers. It requires its two input graphs to be in the PENMAN format for directed graphs (Kasper 1989). The graphs are transformed into triples, and then a mapping between triples from one graph to another is found that maximises the score of overlap. The score is a number between 0 (no overlap at all) and 1 (perfect match).

Figure 10 shows the distributions of the Precision, Recall, and F-Score metrics on the `test` dataset, together with the Macro Average F-Score, which is calculated by collating the entire `test` dataset in a single file, for input into `Smatch`. The baseline at 0.4 represents the average value for the metrics when randomly selecting a pair of images to compare. The overall F-score on the `test` set is 0.67, considerably higher than the baseline.

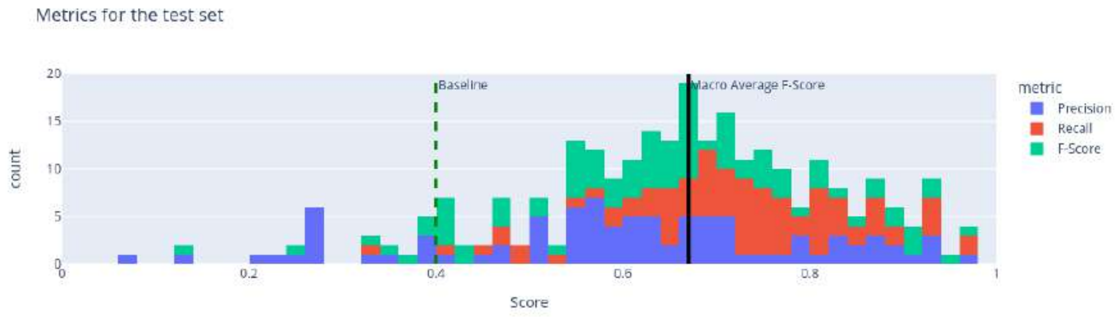


Figure 10: F-scores of Smatch on the `test` dataset.

The averages for the Precision and Recall metrics are between 60% and 75%, with similar results for the `train` and `valid` datasets. This shows that our pipeline generalizes well to images is has not been trained on. It is also clear that there is room for improvement, particularly in the *Semantic Interpretation* component of the pipeline, which integrates and interprets the information provided by the first components. We will discuss our proposed improvements in the next section.

7. Conclusion and Future Work

We showed that it is possible to implement a semantic parser for interpreting tombstone inscriptions by fine-tuning off-the-shelf image object detection and optical character recognition models. The resulting system is far from perfect, but outperforms a baseline system by a wide margin and can therefore be regarded as a successful proof-of-concept system. Indeed there is a lot of space for improvement, which is encouraging for future researchers working in this area.[9]

We see several improvements of the current systems and extensions of the research. Some of these are relatively simple extensions such as interpreting the gender of the deceased, and the interpretation of locative pronouns. Other improvements are more substantial. One of these is the interpretation of biblical references found on tombstones. Although the label detection component is able to predict these reliably, we haven't integrated these in the other components of the pipeline. Another phenomenon that we haven't paid attention to is the interpretation of family roles and relations between people mentioned on tombstone inscriptions. This is a complex problem with more than twenty different types of roles (husband, wife, father, mother, widow, widower, daughter, son, grandfather, grandmother, sister, brother, uncle, aunt, and many more) that require new annotation work for label detection on the image and likewise extensions of the other components of the tombstone parser.

Another extension is the interpretation of toponyms, which is currently done via the names of villages towns, and cities. As there are sometimes spelling variations for the same location, and location names can be ambiguous, even in the same local area (Bos 2022), an interesting alternative is to perform toponym grounding as proposed by Leidner (2008), for instance with the unique identifiers provided by the GeoNames geographical database.[10]

The ultimate type of evaluation is to verify whether the tombstone parser is a significant aid for human curators. Even though it will be unlikely that tombstone parsers will perform on a human

9. The datasets and software will be made available via `http://www.let.rug.nl/bos/tombreader/`.
10. See `https://www.geonames.org/`.

level, it is likely that they can speed up the annotation process. But by how much? A carefully designed experiment with human subjects could answer this question.

References

Artzi, Yoav, Kenton Lee, and Luke Zettlemoyer (2015), Broad-coverage CCG semantic parsing with AMR, *Proceedings of the 2015 Conference on Empirical Methods in Natural Language Processing*, Association for Computational Linguistics, Lisbon, Portugal, pp. 1699–1710.

Banarescu, Laura, Claire Bonial, Shu Cai, Madalina Georgescu, Kira Griffitt, Ulf Hermjakob, Kevin Knight, Philipp Koehn, Martha Palmer, and Nathan Schneider (2013), Abstract Meaning Representation for Sembanking, *Proceedings of the 7th Linguistic Annotation Workshop and Interoperability with Discourse*, Sofia, Bulgaria, pp. 178–186.

Bateman, John A. (1990), Finding translation equivalents: an application of grammatical metaphor, *Proceedings of the 13th conference on Computational Linguistics (COLING 1990)*, pp. 13–18.

Bateman, John A. and Cecile L. Paris (1989), Phrasing a text in terms the user understand, *Proceedings of IJCAI*, pp. 1511–1517.

Bateman, John A., Robert T. Kasper, Joerg F. L. Schuetz, and Erich H. Steiner (1989), A new view on the process of translation, *Proceedings of the fourth conference on European chapter of the Association for Computational Linguistics*, pp. 282–290.

Blackburn, Patrick and Johan Bos (2003), Computational semantics, *Theoria* **18** (46), pp. 27–45.

Bos, Johan (2022), A semantically annotated corpus of tombstone inscriptions, *International Journal of Digital Humanities* **3** (1), pp. 1–33.

Cai, Shu and Kevin Knight (2013), Smatch: an evaluation metric for semantic feature structures, *Proceedings of the 51st Annual Meeting of the Association for Computational Linguistics (Volume 2: Short Papers)*, Association for Computational Linguistics, Sofia, Bulgaria, pp. 748–752.

Eckert, E. (1993), Language change: The testimony of Czech tombstone inscriptions in Praha, Texas, *Varieties of Czech: Studies in Czech sociolinguistics.*

Fellbaum, Christiane, editor (1998), *WordNet. An Electronic Lexical Database*, The MIT Press.

Franken, Morris and Jan C. van Gemert (2013), Automatic Egyptian hieroglyph recognition by retrieving images as texts, *Proceedings of the 21st ACM International Conference on Multimedia*, pp. 765–768.

Howard, A., M. Sandler, B. Chen, W. Wang, L. Chen, M. Tan, G. Chu, V. Vasudevan, Y. Zhu, R. Pang, H. Adam, and Q. Le (2019), Searching for mobilenetv3, *2019 IEEE/CVF International Conference on Computer Vision (ICCV)*, IEEE Computer Society, Los Alamitos, CA, USA, pp. 1314–1324. https://doi.ieeecomputersociety.org/10.1109/ICCV.2019.00140.

Jocher, Glenn, Ayush Chaurasia, Alex Stoken, Jirka Borovec, NanoCode012, Yonghye Kwon, TaoXie, Kalen Michael, Jiacong Fang, imyhxy, Lorna, Colin Wong, Zeng Yifu, Abhiram V, Diego Montes, Zhiqiang Wang, Cristi Fati, Jebastin Nadar, Laughing, UnglvKitDe, tkianai, yxNONG, Piotr Skalski, Adam Hogan, Max Strobel, Mrinal Jain, Lorenzo Mammana, and xylieong (2022), ultralytics/yolov5: v6.2 - YOLOv5 Classification Models, Apple M1, Reproducibility, ClearML and Deci.ai integrations. https://doi.org/10.5281/zenodo.7002879.

Kasper, Robert T. (1989), A flexible interface for linking applications to Penman's sentence generator, *Proceedings of the DARPA Speech and Natural Language Workshop*, Philadelphia, pp. 153–158.

Leidner, Jochen L. (2008), *Toponym Resolution in Text : Annotation, Evaluation and Applications of Spatial Grounding of Place Names*, Universal Press, Boca Raton, FL, USA.

Long, Chris Evin (2016), Disambiguating the departed. Using the genealogist's tools to uniquely identify the long dead and little known, *LRTS* **60** (4), pp. 236–247.

Lyu, Chunchuan and Ivan Titov (2018), AMR parsing as graph prediction with latent alignment, *arXiv preprint arXiv:1805.05286.*

Matthias, George F. (1967), Weathering rates of Portland Arkose tombstones, *Journal of Geological Education* **15** (4), pp. 140–144, Routledge.

Nguyen, Thi Tuyet Hai, Adam Jatowt, Mickael Coustaty, and Antoine Doucet (2021), Survey of Post-OCR processing approaches, *ACM Comput. Surv.*, Association for Computing Machinery, New York, NY, USA. https://doi.org/10.1145/3453476.

Saller, R. and B. Shaw (1984), Tombstones and Roman family relations in the principate: Civilians, soldiers and slaves, *Journal of Roman Studies* **74**, pp. 124–156.

Shi, Baoguang, Xiang Bai, and Cong Yao (2017), An end-to-end trainable neural network for image-based sequence recognition and its application to scene text recognition, *IEEE Transactions on Pattern Analysis and Machine Intelligence* **39** (11), pp. 2298–2304.

Streiter, Oliver and Yoann Goudin (2011), Bury your past, shovel it under: Histories and caterpillars on Taiwan's graveyards, *First Workshop on Documenting and Researching Gravesites in Taiwan*, pp. 27–48.

Streiter, Oliver, Leonhard Voltmer, and Yoann Goudin (2007), From tombstones to corpora: TSML for research on language, culture, identity and gender differences, *Proceedings of the 21st Pacific Asia Conference on Language, Information and Computation (PACLIC)*, Seoul National University, Seoul, Korea, pp. 450–458.

Van Noord, Rik and Johan Bos (2017), Neural semantic parsing by character-based translation: Experiments with abstract meaning representations, *Computational Linguistics in the Netherlands Journal* **7**, pp. 93–108.

Veit, Richard F. and Mark Nonestied (2008), *New Jersey Cemeteries and Tombstones: History in the Landscape*, Rutgers University Press.