IAES International Journal of Robotics and Automation (IJRA)
Vol. 15, No. 1, March 2026, pp. 107~121
ISSN: 2722-2586, DOI: 10.11591/ijra.v15i1.pp107-121

Autonomous reconstruction of strip-shredded documents via self-supervised deep learning and global optimization

Yi-Chang Wu, Pei-Shan Chiang, Yao-Cheng Liu
Department Forensic Science Division, Investigation Bureau, Ministry of Justice, Taipei, Taiwan

Article Info

Article history:

Received Jul 7, 2025
Revised Feb 13, 2026
Accepted Feb 21, 2026

Keywords:

Autonomous reconstruction
Chinese text processing
Forensic science
Fully convolutional neural networks
Global optimization
Self-supervised learning
Strip-shredded documents

ABSTRACT

Autonomous reconstruction of mechanically shredded documents is a labor-intensive challenge in forensic and archival workflows, particularly for scripts with complex structures such as Simplified Chinese. While traditional manual reassembly is tedious, existing digital tools typically rely on extensive human intervention. This paper presents an automated reassembly framework that integrates a lightweight convolutional feature extractor with global combinatorial optimization. By adapting the established SqueezeNet v1.1 backbone, we employ a task-specific self-supervised learning strategy trained on synthetically shredded samples, enabling the adapted model to capture local stroke continuity and edge-geometry cues without manual annotation. The framework infers pairwise relationships from calibrated edge-region inputs, organizing compatibility scores into an asymmetric traveling salesman problem (ATSP) formulation. The optimal fragment sequence is solved deterministically using the Concorde TSP solver, yielding a globally consistent reconstruction. Experimental results on physically shredded documents demonstrate reconstruction accuracies of 86.5% for Simplified Chinese and 94.8% for Western scripts. These results indicate that the proposed pipeline effectively generalizes from synthetic training data to real-world scenarios, providing a practical, high-throughput foundation for automated document recovery under computational constraints typical of robotic or embedded systems.

Corresponding Author:

Yi Chang Wu
Forensic Science Division, Investigation Bureau, Ministry of Justice
No. 74, Zhonghua Rd., Xindian Dist., New Taipei City 231, Taiwan (R.O.C.)
Email: shintenwu@gmail.com

1. INTRODUCTION

The reconstruction of shredded paper documents is a topic of significant relevance to forensic science, investigative disciplines, and archaeology, and has garnered increasing attention in recent years. Following the fall of the Berlin Wall in 1989, the Ministry for State Security of East Germany (Stasi) attempted to destroy a vast amount of intelligence documents, resulting in over 16,000 bags of shredded materials. As shown in Figure 1 [1], an employee is holding fragments of the shredded Stasi files. The German government mobilized approximately 40 personnel, and after six years of effort, only 300 documents were successfully reconstructed [2]. The recovery process remains ongoing to this day. Schauer *et al.* [3] suggested that shredded documents can be regarded as a variation of conventional jigsaw puzzles. Puzzle-solving techniques have been extensively applied in diverse domains, including biosensing [4]–[6], image reconstruction [7]–[10], speech unscrambling [11], [12], banking [13], and forensic analysis [14]–[16]. In archaeology, such techniques assist in the identification of cultural heritage and the restoration of artifacts [17]–[20]. In legal investigations, they are employed to recover deliberately damaged documents and

photographs [21]–[23]. Furthermore, in military contexts—particularly in combat zones—they are crucial for recognizing remnants of destroyed documents. In summary, fragment reconstruction technologies play a vital role in the retrieval of forensic evidence, preservation of cultural assets, and acquisition of military intelligence.

Due to the irregular shapes, diverse sizes, and large number of remaining fragments, manual reconstruction is highly inefficient and often infeasible. Consequently, the complete reconstruction process, encompassing scanning, analysis, and reassembly, must be amenable to full automation. The methodology presented here is designed as a core component within an automated pipeline. The integration begins at the document digitization stage, where high-resolution overhead scanners capture fragment images in a consistent and repeatable manner. These digitized fragments are then directly streamed into the proposed deep learning-based reconstruction framework, which performs compatibility estimation and global reassembly without human intervention. By combining physical document sensing with autonomous visual perception and optimization-based decision making, the proposed system naturally fits into robotic and automated document processing workflows.

Figure 1. A staff member displaying shredded documents originally destroyed by the Stasi, an organization known for its extensive informant network used to monitor East German citizens [1]

The efficiency of manual reconstruction of shredded documents is influenced by several factors, including the complexity of the document, the nature of the destruction process, and the number and shape of the fragments. Even with the aid of artificial intelligence systems, these factors significantly impact the time required for reconstruction. Research on shredded document reconstruction encompasses multiple subfields, which are often differentiated by shredding methods (e.g., mechanical cutting or manual tearing), document types (e.g., black-and-white or color, text-only, images, or mixed content), and reconstruction approaches (e.g., fully automated or human-assisted systems). Currently, most studies leveraging modern technology for reconstruction focus on simulating shredding scenarios, where identifying and integrating adjacent fragments remains the core challenge. Given that most documents are destroyed via mechanical shredding rather than manual tearing, and considering that black-and-white text documents are the most common type, this study focuses on developing a reconstruction system for mechanically shredded black-and-white printed documents. Additionally, we center our work on Simplified Chinese—a widely used yet under-explored language in the literature.

Reconstruction research fundamentally relies on visual cues. This presents a particular challenge for text-based documents due to their limited color information (i.e., binary black-and-white appearance). Early approaches to text document reconstruction primarily relied on puzzle-piece shape features to determine fragment adjacency [24]–[26]. However, since mechanically shredded fragments tend to have highly similar shapes, conventional shape-based algorithms—such as those using polygonal approximation to simplify curve matching—are unsuitable for our target scenario. Some methods [27]–[31] have focused on color distribution, utilizing color information to assess the compatibility between fragments. While such approaches leverage the richer information present in color images compared to binary (black-and-white) data, they typically neglect the edge degradation caused by mechanical shredding. As a result, they are not applicable to the reconstruction of black-and-white textual documents. Lin *et al.* [32] introduced an approach using average word length to represent English characters and applied fragment encoding to describe document layouts. Prandtstetter and Raidl [33] formulated the reconstruction of strip-shredded documents as

a variant of the classical Traveling Salesman Problem (TSP), and proposed a variable neighborhood search method to optimize the reconstruction process in a semi-automated framework. Balme [34] and Morandell [35] employed binary image representations to model the black-and-white appearance of textual documents. They addressed the issue of vertical misalignment between adjacent fragments by respectively computing weighted pixel correlation and quantifying the degree of misalignment between black pixel regions. Li *et al.* [36] further advanced these rule-based approaches by using geometric templates and blank-area detection to reconstruct rectangular English fragments. However, such methods depend heavily on the predictability of Western character layouts and often struggle with the irregular stroke densities of other scripts. Some studies have evaluated fragment compatibility by analyzing the degree of character continuity along shredded edges. For instance, Perl *et al.* [37] investigated the use of OCR features for supervised character recognition and alignment, proposing an English OCR-based method that employs character histograms to match fragment boundaries. Their findings indicated that reconstruction accuracy decreases non-monotonically as the number of text lines diminishes. Paixao *et al.* [38] analyzed the shapes of character groupings to classify different symbol combinations and calculate the compatibility of fragment pairs.

Deep learning has achieved state-of-the-art performance in computer vision tasks such as image classification, object detection, and segmentation. Unlike earlier template-matching techniques [36], convolutional neural networks (CNNs) are capable of capturing fine-grained stroke continuity and learning task-specific representations directly from raw pixels. Sholomon *et al.* [39] used neural networks to predict whether two puzzle piece edges should be adjacent by feeding their edge pixel information into the network. However, these methods were designed for synthetically generated fragments and not for real-world shredded documents. Most prior research focuses on Western languages, which differ significantly from Chinese in terms of character structure. Chinese characters are square-shaped, spatially uniform, and independent units, unlike Western languages which are composed of linear, horizontally arranged letters. Standard printed Chinese characters exhibit a 1:1 height-to-width ratio. When Chinese documents are shredded, the resulting fragments may contain either damaged characters along the edges or blank regions corresponding to interline spacing. This study attempts to address the reconstruction of shredded Chinese documents under real-world conditions by leveraging deep learning models [40] in a self-supervised learning framework, enabling large-scale sample extraction and learning from unlabeled data. Our findings may also offer valuable insights for document reconstruction in languages with similar logographic writing systems, such as Japanese and Korean.

The remainder of this paper is organized as follows. Section 2 details the proposed autonomous reconstruction framework, encompassing document digitization, its integration with automated scanning systems, self-supervised sample generation, model training with the SqueezeNet backbone, and the global optimization search via ATSP. Section 3 reports the experimental results, including a comprehensive performance benchmarking against existing methodologies. Section 4 concludes the paper by summarizing our findings and discussing potential future directions for real-world robotic applications.

2. METHOD

This study aims to develop a model that quantifies the compatibility between pairs of document fragments. Due to the labor-intensive nature of creating real-world shredded datasets and the lack of public datasets for strip-shredded documents, we adopt a self-supervised learning approach that automatically determines fragment adjacency during the sampling process. Specifically, we simulate document shredding digitally and extract fragment pairs as training samples. In this process, adjacent fragment pairs are labeled as positive samples, while non-adjacent pairs are labeled as negative samples. A fully convolutional neural network (FCNN) is trained as a binary classifier. The best-performing model is used to evaluate pairwise compatibility based on local visual content of each fragment. The resulting matching scores are stored in a matrix, which is subsequently used as input for a graph-based optimization algorithm to determine the optimal reassembly sequence.

The model is validated using real shredded documents. The following sections detail the key steps of the proposed system, including document digitization, training sample generation, self-supervised training, pairwise compatibility scoring, and optimization-based reassembly. The overall workflow is illustrated in Figure 2.

2.1. Document digitization

Commercial shredders typically produce either strip-cut (spaghetti-like) or cross-cut fragments. In this study, we focus on the reconstruction of strip-shredded documents, which represent the most common shredding mechanism used in practical forensic and archival scenarios. The digitization process serves as the entry point of the proposed reconstruction pipeline and is designed to support automated, high-throughput processing.

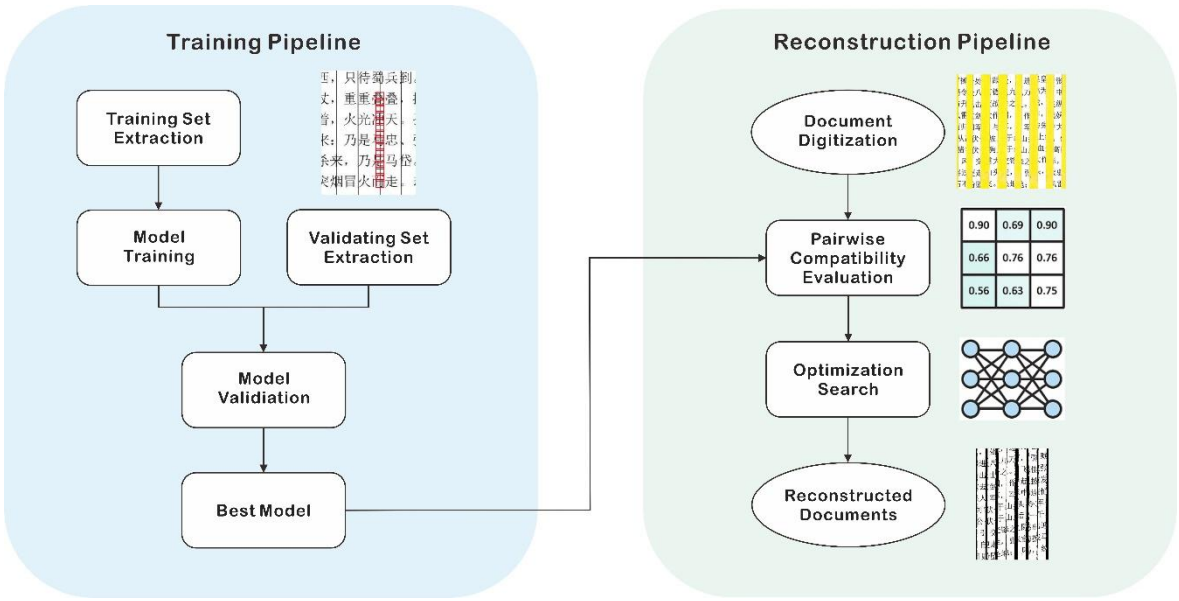


Figure 2. System workflow for shredded document reconstruction

Printed documents are first mechanically shredded, and visually blank fragments are discarded. The remaining fragments are mounted on a high-saturation, non-grayscale background to facilitate robust foreground–background separation. This design choice enables reliable automated processing in downstream stages without requiring manual annotation or intervention.

Fragment extraction is performed using k-means clustering in the RGB color space, where image pixels are grouped into three classes corresponding to fragment content, paper substrate, and background. After identifying the background cluster, all associated pixels are removed to obtain clean fragment contours. Each fragment is then isolated and stored as an individual image. Binarization is subsequently applied using the Sauvola method [41]. Since fragments are placed with consistent orientation during acquisition, no rotation correction is required, thereby reducing geometric distortion and preserving edge information critical for compatibility estimation. All extracted fragments are standardized and forwarded to the learning-based reconstruction stage.

2.1.1. Integration with automated scanning systems

The document digitization procedure is inherently compatible with automated scanning and robotic document-processing pipelines. Overhead scanners, such as the ScanSnap SV600 employed in this study, provide non-contact image acquisition and stable imaging geometry, making them well suited for integration into automated forensic or archival systems.

In an operational setting, shredded fragments can be sequentially scanned without manual alignment, after which all subsequent stages—including background removal, contour extraction, binarization, fragment compatibility estimation, and optimization-based reassembly—are executed entirely in software on a processing unit directly connected to the scanner. This end-to-end automation establishes a seamless data flow from physical fragment acquisition to digital reconstruction, positioning the proposed framework as a core computational component within automated document recovery and robotic information-processing systems.

2.2. Training sample preparation

To construct the training dataset, we prepared an internally generated corpus consisting of 300 pages of Simplified Chinese documents and 100 pages of English documents, all digitally created in standard office formats. All documents were converted into 300-dpi grayscale images prior to preprocessing. To generate sufficient training data for the boundary classifier, each document was digitally shredded, and 50% of the documents were allocated to training while the remaining 50% were reserved for validation, ensuring that no document contributed samples to both sets.

Each document image was first binarized using the Sauvola method [41] and then partitioned into 30 vertical strips of equal width, each preserving the full height of the original image. To simulate the

irregularities commonly observed in physical shredded edges and avoid overly smooth strip boundaries, the outermost two-pixel columns on both sides of each strip were replaced with pseudo-random black-and-white patterns drawn from a uniform binary distribution U(0,1).

Before sample extraction, the simulated strips from each document were randomly shuffled to minimize sampling bias and increase coverage across different textual regions. Training samples were extracted along the boundary between each pair of adjacent strips. Specifically, a 32×32 patch was cropped every two pixels, consisting of 16 pixels from each side of the boundary. Patches generated from correctly matched strip pairs were labeled as positive, whereas those from mismatched pairs were labeled as negative. An equal number of positive and negative samples were collected per document.

Across the 400-page corpus, this process produced approximately 14 million candidate patches. To ensure that only informative boundary regions were retained, samples with a foreground-pixel ratio below 0.1 were discarded, as such patches typically correspond to blank margins or scanning noise. The resulting dataset—stored in binary image format and assuming black text on a white background—contained roughly 6 to 7 million samples each for the training and validation sets.

2.3. Feature extraction backbone and computational rationale

We selected SqueezeNet v1.1 [42], pretrained on ImageNet, as the backbone feature extractor for fragment compatibility prediction. SqueezeNet v1.1 is a fully convolutional neural network designed to achieve competitive recognition performance under a restricted parameter and memory budget. A primary design consideration in our framework is the balance between representational capacity and operational throughput. While high-capacity architectures such as ResNet or EfficientNet offer increased depth and improved performance on global semantic benchmarks, their computational overhead is often prohibitive for high-throughput robotic reassembly pipelines.

In the present task, fragment compatibility is determined primarily by local stroke continuity and edge-geometry alignment rather than high-level semantic abstraction. Consequently, the effective feature space is significantly less complex than that required for general object recognition. Deeper architectures therefore provide diminishing returns for boundary matching while introducing substantial latency and memory overhead. Furthermore, practical document reassembly requires evaluating thousands of candidate fragment pairings, resulting in an inherent $O(n^2)$ computational complexity. Under this constraint, a lightweight backbone is essential to prevent the vision module from dominating system latency.

Accordingly, SqueezeNet v1.1 is specifically tailored to these requirements, offering representational capability comparable to larger deep networks while utilizing approximately 50 times fewer parameters. This reduction in model complexity is instrumental for deployment in resource-constrained or latency-sensitive environments, ensuring that the reconstruction pipeline remains viable for real-time automated forensic workflows. As illustrated in Figure 3, the network begins with an initial convolutional layer (conv1), followed by eight Fire modules (Fire2–Fire9) interleaved with three max-pooling layers, and terminates with a global average pooling layer. The internal structure of each Fire module is shown in Figure 4 and consists of a squeeze layer with 1×1 filters and an expand layer combining parallel 1×1 and 3×3 filters. This architectural design minimizes parameter count while preserving the fine-grained textual features required for accurate fragment alignment. In this study, the original SqueezeNet v1.1 configuration is retained without structural modification to ensure reproducibility and compatibility with real-time automated document reconstruction workflows.

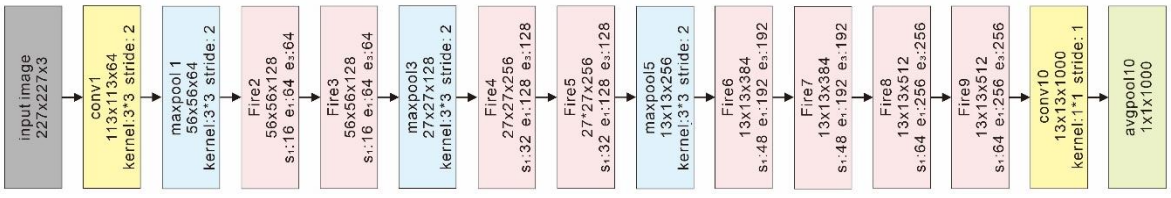


Figure 3. The architecture of SqueezeNet 1.1

2.4. Self-supervised training and pairwise compatibility scoring

To adapt the ImageNet-pretrained backbone to the document reconstruction task, each binary fragment-pair image is replicated across three channels to form a fixed-size 227×227×3 input. The final convolutional layer is replaced with a two-filter output corresponding to binary classification (compatible versus incompatible), with weights initialized from a zero-mean Gaussian distribution with a standard

deviation of 0.01. Training is performed for 10 epochs using the Adam optimizer [43] and categorical cross-entropy loss, with a mini-batch size of 256. Model performance is evaluated on a validation set at the end of each epoch, and the checkpoint achieving the highest validation accuracy is selected for subsequent inference.

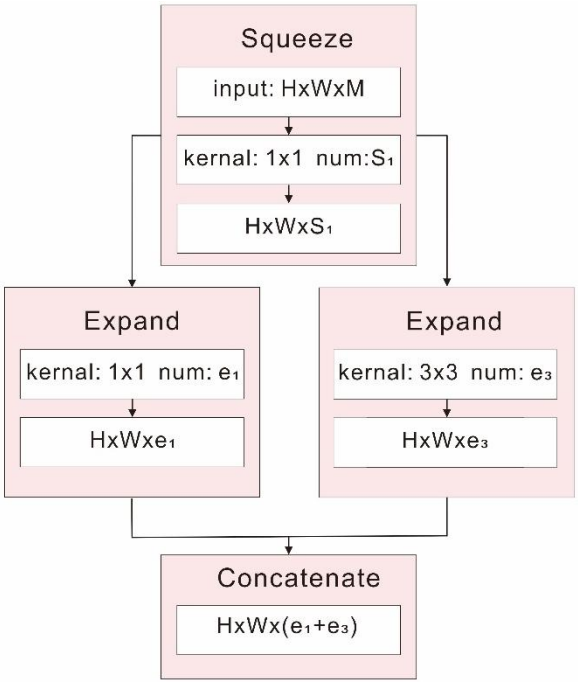


Figure 4. The architecture of fire module

During inference, the trained network evaluates the pairwise compatibility of non-blank fragments $F = \{f_1, f_2, \ldots, f_n\}$. For each ordered pair (f_p, f_q), a likelihood score M_{pq} is computed to estimate the probability that f_q is the immediate right neighbor of f_p. Each evaluation uses a calibrated image of size H×32, composed of the rightmost 16 pixels of f_p and the leftmost 16 pixels of f_q. To compensate for vertical misalignment commonly introduced during mechanical shredding, a vertical offset parameter m=10 is applied, resulting in 21 candidate evaluations per fragment pair. The maximum probability among these candidates is retained as the final compatibility score in matrix M.

2.5. Global optimization via ATSP formulation

The optimal fragment sequence is obtained by formulating the reassembly task as Asymmetric Traveling Salesman Problem (ATSP). A distance matrix N is derived from the compatibility matrix M, where N_{pq}= max(M) $-M_{pq}$ for $p \neq q$., and diagonal elements are set to infinity. This formulation defines a directed weighted graph in which each vertex corresponds to a fragment.

To find the optimal global sequence, the problem is treated as finding the shortest path that visits each node exactly once. This is achieved by introducing a virtual node connected to all fragments via zero-weight edges, effectively transforming the fragment ordering task into a standard ATSP. To leverage the industry-standard Concorde TSP solver [44], the ATSP is further converted into a symmetric TSP using the two-node transformation method [45] and solved exactly with the QSOpt3 library. The resulting fragment ordering is deterministic and globally optimized, providing a reliable high-level execution reference for automated or robotic document reassembly systems.

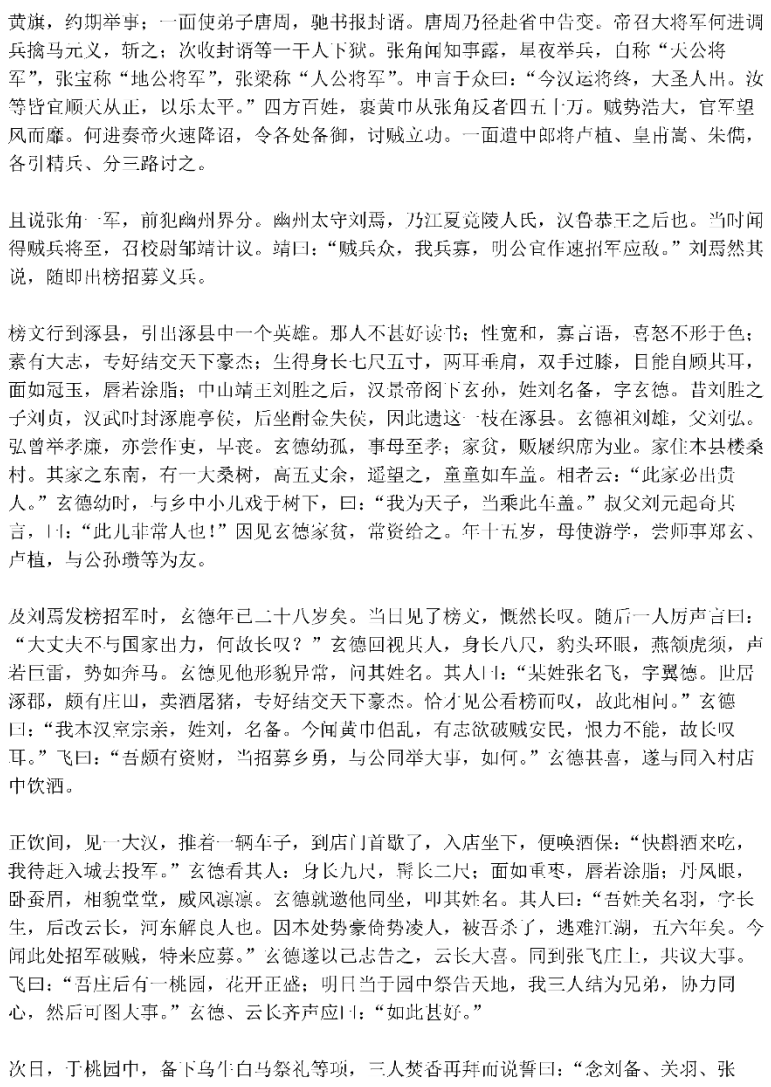
3. RESULTS AND DISCUSSION

3.1. Experimental datasets and preprocessing

To evaluate the feasibility and alignment accuracy of the proposed shredded document reconstruction method on Simplified Chinese texts and to explore its performance on Western languages, we conducted experiments on two real-world shredded datasets: the D2-mec dataset [46] and the Csim dataset. The D2-mec dataset consists of fragments from 20 English plain-text documents sourced from the ISRI-Tk

OCR database [47], which were shredded using a Leadership model 7348 shredder. The fragments were manually reassembled, scanned at a resolution of 300 dpi, and digitized into image format.

The Csim dataset was specifically constructed to address the lack of publicly available datasets involving the document types targeted in this study. An example is shown in Figure 5, documents were generated in Microsoft Word using A4 paper, SimSun font, left-to-right horizontal text layout, single-line spacing, and a font size of 12 pt. After shredding with an IDEAL 2260 strip-cut shredder (4 mm width), non-informative blank fragments were removed. The fragment edges were irregular and occasionally exhibited minor information loss, reflecting the physical characteristics of real shredded documents, in contrast to the uniformity seen in digitally simulated fragments. The remaining fragments—approximately 38–41 per page—were affixed to uniformly saturated yellow A4 sheets and digitized using a ScanSnap SV600 overhead scanner at 300 dpi. Fragments were then extracted from the scanned image and saved as individual files.



黄旗，约期举事；一面使弟子唐周，驰书报封谞。唐周乃径赴省中告变。帝召大将军何进调兵擒马元义，斩之；次收封谞等一干人下狱。张角闻知事露，星夜举兵，自称“天公将军”，张宝称“地公将军”，张梁称“人公将军”。申言于众曰：“今汉运将终，大圣人出。汝等皆宜顺天从正，以乐太平。”四方百姓，裹黄巾从张角反者四五十万。贼势浩大，官军望风而靡。何进奏帝火速降诏，令各处备御，讨贼立功。一面遣中郎将卢植、皇甫嵩、朱儁，各引精兵、分三路讨之。

且说张角一军，前犯幽州界分。幽州太守刘焉，乃江夏竟陵人氏，汉鲁恭王之后也。当时闻得贼兵将至，召校尉邹靖计议。靖曰：“贼兵众，我兵寡，明公宜作速招军应敌。”刘焉然其说，随即出榜招募义兵。

榜文行到涿县，引出涿县中一个英雄。那人不甚好读书；性宽和，寡言语，喜怒不形于色；素有大志，专好结交天下豪杰；生得身长七尺五寸，两耳垂肩，双手过膝，目能自顾其耳，面如冠玉，唇若涂脂；中山靖王刘胜之后，汉景帝阁下玄孙，姓刘名备，字玄德。昔刘胜之子刘贞，汉武时封涿鹿亭侯，后坐酎金失侯，因此遗这一枝在涿县。玄德祖刘雄，父刘弘。弘曾举孝廉，亦尝作吏，早丧。玄德幼孤，事母至孝；家贫，贩屦织席为业。家住本县楼桑村。其家之东南，有一大桑树，高五丈余，遥望之，童童如车盖。相者云：“此家必出贵人。”玄德幼时，与乡中小儿戏于树下，曰：“我为天子，当乘此车盖。”叔父刘元起奇其言，曰：“此儿非常人也！”因见玄德家贫，常资给之。年十五岁，母使游学，尝师事郑玄、卢植，与公孙瓒等为友。

及刘焉发榜招军时，玄德年已二十八岁矣。当日见了榜文，慨然长叹。随后一人厉声言曰：“大丈夫不与国家出力，何故长叹？”玄德回视其人，身长八尺，豹头环眼，燕颔虎须，声若巨雷，势如奔马。玄德见他形貌异常，问其姓名。其人曰：“某姓张名飞，字翼德。世居涿郡，颇有庄山，卖酒屠猪，专好结交天下豪杰。恰才见公看榜而叹，故此相问。”玄德曰：“我本汉室宗亲，姓刘，名备。今闻黄巾倡乱，有志欲破贼安民，恨力不能，故长叹耳。”飞曰：“吾颇有资财，当招募乡勇，与公同举大事，如何。”玄德甚喜，遂与同入村店中饮酒。

正饮间，见一大汉，推着一辆车子，到店门首歇了，入店坐下，便唤酒保：“快斟酒来吃，我待赶入城去投军。”玄德看其人：身长九尺，髯长二尺；面如重枣，唇若涂脂；丹凤眼，卧蚕眉，相貌堂堂，威风凛凛。玄德就邀他同坐，叩其姓名。其人曰：“吾姓关名羽，字长生，后改云长，河东解良人也。因本处势豪倚势凌人，被吾杀了，逃难江湖，五六年矣。今闻此处招军破贼，特来应募。”玄德遂以己志告之，云长大喜。同到张飞庄上，共议大事。飞曰：“吾庄后有一桃园，花开正盛；明日当于园中祭告天地，我三人结为兄弟，协力同心，然后可图大事。”玄德、云长齐声应曰：“如此甚好。”

次日，于桃园中，备下乌牛白马祭礼等项，三人焚香再拜而说誓曰：“念刘备、关羽、张

Figure 5. Sample document used for reconstruction

Figure 6 illustrates this preprocessing pipeline. Figure 6(a) shows the original fragments pasted onto yellow background paper. Figure 6(b) presents 10 sample fragments extracted from 6(a), which were binarized such that the fragments appear white, the background black, and the character strokes in black, enabling effective foreground-background separation.

Table 1 summarizes the total number of fragments and the number of correctly assembled fragments in the test documents from the D2-mec and Csim datasets. For the 20 documents in the D2-mec dataset, each document contains between 20 and 28 fragments, with 0 to 2 misassembled fragments per case. The Csim

dataset contains substantially more fragments (38–41 per page), with 3–7 assembly errors per document. Figure 7 presents examples of reconstruction results, where Figures 7(a) and 7(b) are drawn from the D2-mec and Csim datasets, respectively. An examination of the reconstructed outputs revealed that the fragments were placed side-by-side without rotation, and the vertical alignment difference between fragments was minimized.

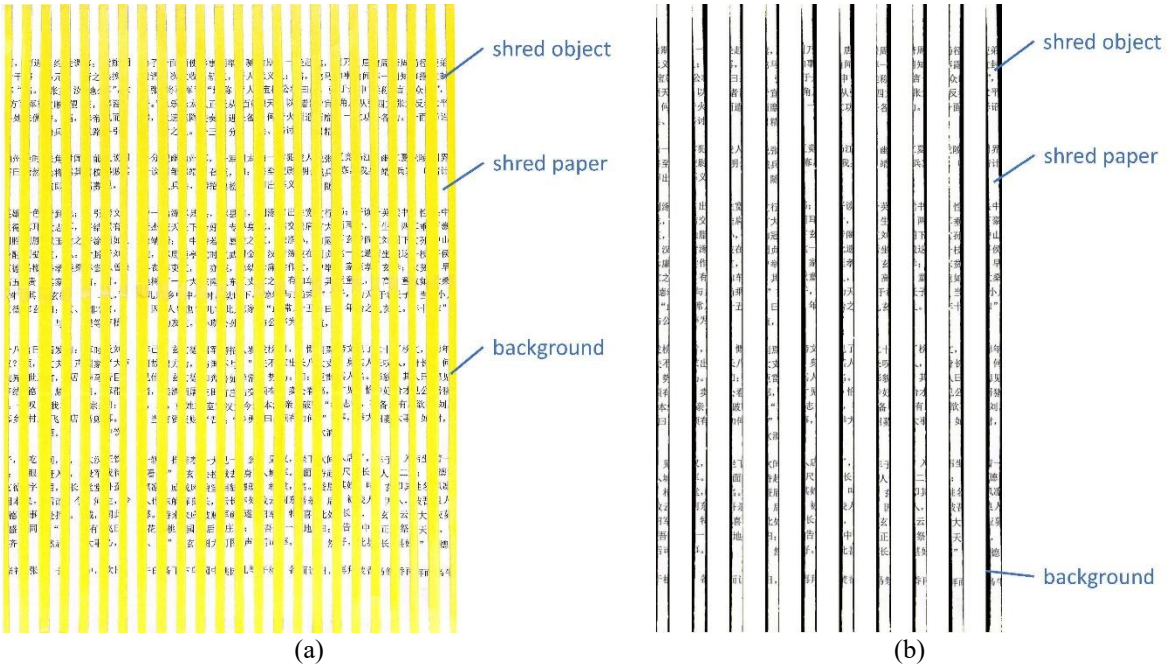
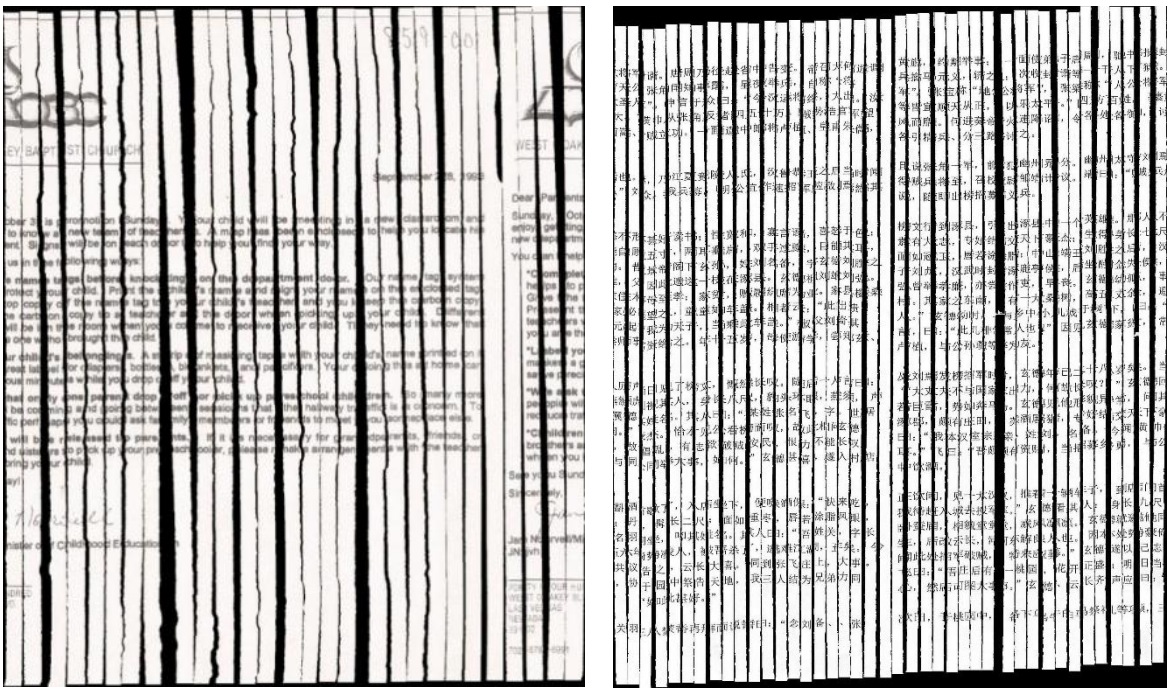


Figure 6. Preprocessing results of shredded paper fragments (a) fragments pasted on yellow background paper and (b) ten extracted and binarized fragment images

Table 1. Total number of fragments and number of correctly matched pairs in the D2-mec and Csim datasets

Dataset	D2-mec		Csim	
No.	total	correct	total	correct
1	26	25	40	35
2	28	27	40	34
3	28	27	40	33
4	26	25	40	35
5	24	23	40	34
6	27	26	40	36
7	27	26	41	35
8	25	23	41	37
9	27	26	40	35
10	26	24	40	36
11	25	23	41	35
12	25	24	39	33
13	20	18	41	34
14	25	23	40	34
15	26	25	39	32
16	24	24	39	35
17	25	23	38	34
18	26	25	38	32
19	24	23	40	33
20	23	21	41	35
21			40	36
22			41	38
23			41	35
24			38	34
25			40	33



(a) (b)

Figure 7. Reconstruction results (a) sample from the D2-mec dataset with 96.4% accuracy; (b) Sample from the Csim dataset with 92.7% accuracy

3.2. Reconstruction accuracy and quantitative evaluation

3.2.1. Accuracy definition

To quantify reconstruction quality, we define reconstruction accuracy as:

$$\text{Accuracy} = 1 - \frac{\text{Number of incorrectly positioned fragments}}{\text{Total number of fragments}}$$

3.2.2. Primary results

Figure 8 illustrates the reconstruction accuracy achieved on individual test documents. The results indicate that the per-document reconstruction accuracy exceeds 90.0% for English documents and 82.1% for Chinese documents. When averaged on a document-wise basis, the proposed method achieves an average reconstruction accuracy of 94.8% on the D2-mec dataset and 86.5% on the Csim dataset, respectively.

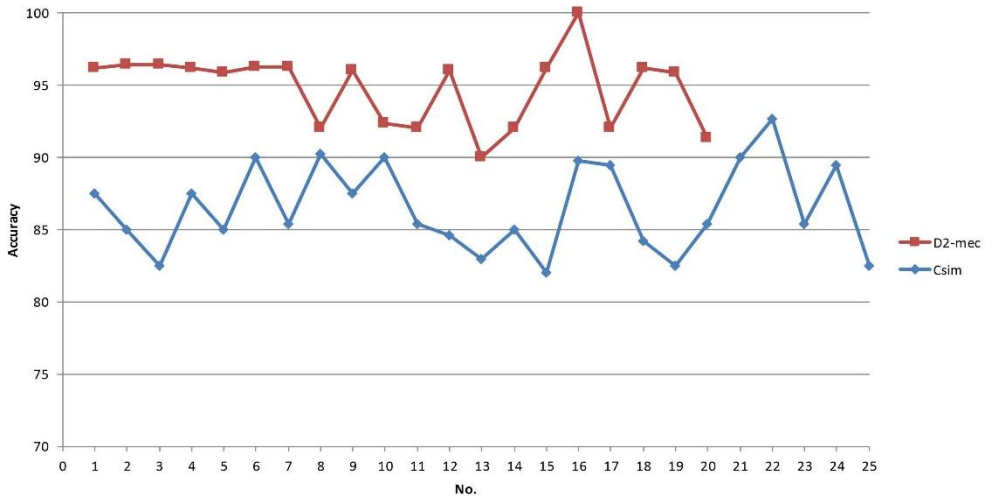


Figure 8. Accuracy of reconstruction results

3.2.3. Reliability analysis

In addition to reconstruction accuracy, precision, recall, and F1-score were adopted to further characterize the reliability of fragment placement. These metrics provide complementary insights into the correctness and completeness of the reconstruction process.

In the proposed setting, each fragment is assigned to a unique position without the generation of spurious or duplicate fragments. Consequently, reconstruction errors arise exclusively from misplacements (swapping) rather than incorrect fragment insertions. Under this constraint, the system achieved a precision of 1.0 for both the D2-mec and Csim datasets, confirming that every placed fragment corresponds to a valid document component.

Table 2 reports the averaged precision, recall, and F1-score across all test documents. The results demonstrate high reconstruction reliability, with the F1-score remaining robust across both English and Chinese scripts. While the recall—which is equivalent to document-level reconstruction accuracy—varies between the two datasets, the consistently high precision and F1-scores underscore the stability of the self-supervised FCNN in extracting discriminative features from physical fragments.

Table 2. Evaluation metrics based on per-document averaging

Dataset	Precision	Recall	F1-score
D2-mec	1	0.948	0.973
Csim	1	0.865	0.928

3.3. Analysis of influencing factors: script structure and stroke complexity

In most cases, the reconstruction accuracy for English text was higher than that for Chinese text. This discrepancy can be attributed to several factors, one of which is the structural characteristics of Chinese characters. Chinese characters are typically square-shaped, uniformly sized, and evenly arranged, which often introduces vertical whitespace columns between character blocks. When the shredder cuts along these whitespace columns—resulting in edge fragments containing minimal information—these non-informative edge fragments may be mistakenly reassembled together during reconstruction. In contrast, characters in English text are generally more irregularly arranged, making it less likely for large blank columns to form, and thus reducing the chances of such reassembly errors.

The number of fragments may also affect reconstruction accuracy. With fewer fragments, there are fewer pairing candidates, which may reduce the probability of mismatches. In the D2-mec dataset, the number of fragments per document was relatively small, whereas the Csim dataset contained significantly more fragments per page. To validate this hypothesis that fewer fragments would increase reconstruction accuracy, we conducted an additional experiment in which several Csim documents were manually torn into eight irregular strip-like fragments. The results showed that the reconstruction accuracy reached 100%. Figure 9 shows an example of one of these reconstructions. This suggests that reconstruction accuracy is indeed influenced by the number of fragments.

We also investigated the influence of stroke density and character complexity. Chinese characters tend to have more complex stroke structures. When shredded, Chinese characters produce a large number of disjointed strokes, resulting in a greater number of candidate matching points during reconstruction. In contrast, Western characters are structurally simpler, with fewer stroke discontinuities. We hypothesize that character sets with higher stroke complexity or smaller font sizes may lead to increased reconstruction errors, as the density of stroke breakpoints per unit area becomes higher. To validate this hypothesis, we increased the font size of Chinese characters in the test documents and conducted reconstruction experiments using 4 mm-wide machine-shredded fragments. The results confirmed that enlarging the font size reduced the density of stroke discontinuities per unit area and significantly improved reconstruction accuracy. As shown in Figure 10, when the font size in the test example (Figure 5) increased from 12 to 28, the reconstruction accuracy reached 100%. These findings suggest that reconstruction accuracy is also influenced by the stroke complexity and font size of the script.

Compared with synthetic fragments used in simulated experiments, real shredded documents introduce several factors that can degrade reconstruction performance. These include edge damage from the shredding process, angular distortions between fragments and the original layout caused by mechanical cutting and digitization, scanner resolution limitations, image noise introduced during scanning, vertical misalignment between adjacent fragments, and contour extraction noise. In addition, the number of fragments, the nature of the script, and font size also play critical roles. Despite these challenges, our experiments demonstrate the feasibility of the proposed reconstruction method on real shredded Chinese documents.

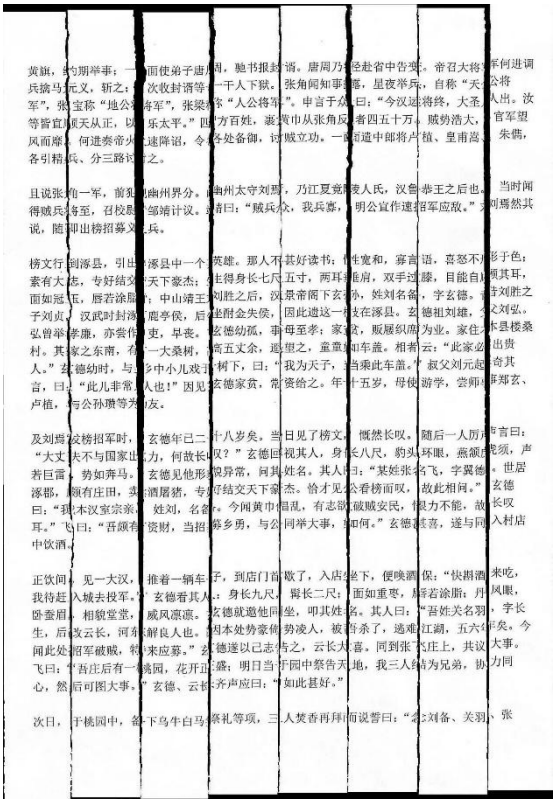


Figure 9. Reconstruction example with manually torn document in Figure 5 into eight long strips, achieving 100% accuracy

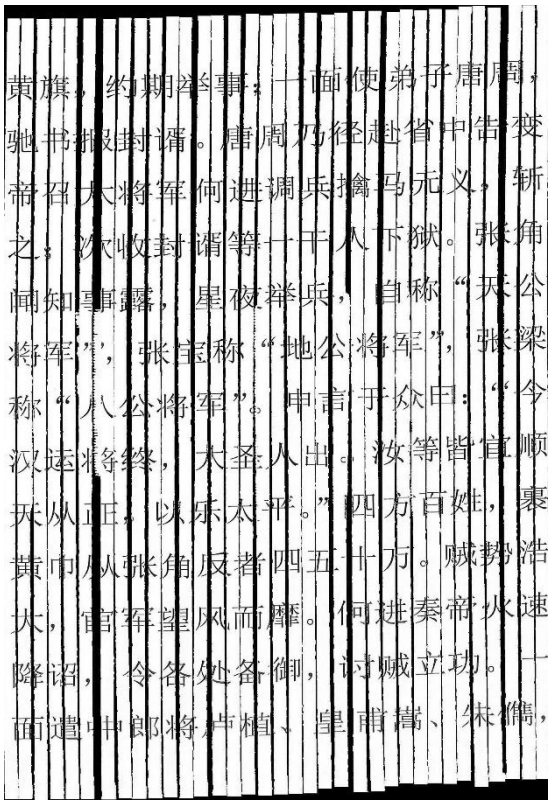


Figure 10. Reconstruction example with enlarged font size (from 12pt to 28pt) in the document shown in Figure 3, achieving 100% accuracy

3.4. Robotic automation and practical integration potential

From an architectural perspective, the proposed reconstruction framework is inherently compatible with automated document handling and robotic systems. By formulating fragment reassembly as an asymmetric traveling salesman problem (ATSP), the system produces a deterministic and globally optimized fragment sequence that serves not only as a digital result but also as a high-level action plan for physical execution.

In a potential robotic deployment, the optimized fragment order generated by the Concorde-based solver can be directly translated into motion sequences for a robotic manipulator equipped with precision end-effectors, such as vacuum suction grippers. This allows fragments to be physically repositioned according to the computed Hamiltonian path while respecting spatial alignment constraints. The modular design of the pipeline—separating acquisition, deep-learning-based compatibility estimation, and global optimization—enables seamless integration with external robotic subsystems. Compared with manual reconstruction, which is labor-intensive and prone to subjective error, the proposed method provides a robust algorithmic decision-making layer essential for end-to-end autonomous physical reconstruction.

3.5. Performance benchmarking and comparative analysis

To validate the robustness of the autonomous decision-making layer, we conducted a focused horizontal comparison against representative methodologies, ranging from classical geometric matching to contemporary deep learning. This analysis highlights a significant transition from manual, feature-engineered approaches toward the fully autonomous, self-supervised pipeline proposed in this study.

Environmental Fidelity and Robustness: A primary distinction of this work is its emphasis on environmental fidelity. Many existing studies, most notably the texture-feature matching approach by Li *et al.* [36] and the minimum-error method by Le *et al.* [48], primarily validated their algorithms using digitally shredded (synthetic) samples. While effective for algorithmic verification, synthetic slicing bypasses the complex physical artifacts—such as edge roughness, fiber protrusion, and mechanical distortion—inherent in real-world forensic scenarios. Our pipeline addresses this gap by achieving a robust 86.5% accuracy on physically shredded Chinese documents. This result demonstrates superior resilience to the stochastic noise of mechanical shredding compared to the pixel-level matching proposed in [36] and [48], proving the efficacy of our semantic feature extraction in realistic forensic environments.

Automation and Scalability: Regarding the level of automation, which is foundational for robotic systems, our framework offers significant advantages over existing semi-automatic tools. The hybrid approaches presented by Prandtstetter and Raidl [33] and De Smet *et al.* [49] rely on user interactions and GUI-based "digital gluing" to resolve alignment ambiguities. While such "human-in-the-loop" systems are effective for small-scale applications, they lack the scalability required for high-throughput robotic reassembly. In contrast, our pipeline achieves full end-to-end autonomy, providing a deterministic decision-making layer that integrates seamlessly with robotic motion planners without human intervention.

Cross-Script Generalization: Furthermore, while Alhaj *et al.* [50] reported a high accuracy of 96.2%, their methodology is predicated on color-gradient consistency, which is typically absent in monochrome textual forensics. Our framework excels in these "information-poor" scenarios by extracting high-level semantic features through self-supervised learning. Finally, our performance on Western scripts (94.8%) is highly consistent with the state-of-the-art Baseline DL model [40] (96.4%), suggesting that our approach expands the applicability of deep-learning-based reconstruction to the structurally complex domain of Simplified Chinese text without compromising performance on alphabetic scripts.

4. CONCLUSION

This study demonstrates the feasibility of reconstructing mechanically shredded documents—specifically complex Simplified Chinese texts—under real-world conditions using a systematic framework built upon self-supervised learning and combinatorial optimization. By incorporating an existing lightweight convolutional neural network as a compatibility estimator, we have shown that high-fidelity reassembly can be achieved while maintaining the robustness and scalability required for practical deployment. The strategic integration of deep-learning-based feature extraction with exact global optimization allows the system to overcome challenges such as edge damage and vertical misalignment, successfully generalizing from large-scale synthetic training samples to physically shredded artifacts.

Despite these results, several limitations warrant further investigation. The system's performance may degrade in "information-poor" scenarios involving mixed-content pages or color documents. Furthermore, while the current pipeline is fully automated in the digital domain, it serves as a high-level decision-making layer that does not yet address physical fragment manipulation.

Looking forward, the deterministic fragment ordering produced by our optimization stage provides a natural interface for robotic handling systems. By translating the computed sequence into motion commands, future work can bridge the gap between digital reconstruction and physical reassembly. This study confirms that leveraging established deep learning models within a carefully structured optimization pipeline offers a practical and extensible solution for real-world forensic recovery in automated and robotic contexts.

ACKNOWLEDGMENTS

The authors would like to express their gratitude to the Ministry of Justice, Taiwan, for providing the necessary resources and financial support for this research.

FUNDING INFORMATION

The authors express their gratitude to the Ministry of Justice of Taiwan for financially supporting this study through the Science and Technology Project (112-1301-10-28-02 and 113-1301-10-28-02).

AUTHOR CONTRIBUTIONS STATEMENT

This journal uses the Contributor Roles Taxonomy (CRediT) to recognize individual author contributions, reduce authorship disputes, and facilitate collaboration.

Name of Author	C	M	So	Va	Fo	I	R	D	O	E	Vi	Su	P	Fu
Yi-Chang Wu	✓	✓	✓	✓			✓			✓		✓	✓	✓
Pei-Shan Chiang		✓		✓	✓	✓		✓	✓		✓			
Yao-Cheng Liu		✓	✓	✓	✓	✓			✓		✓			

C : **C**onceptualization
M : **M**ethodology
So : **So**ftware
Va : **Va**lidation
Fo : **Fo**rmal analysis
I : **I**nvestigation
R : **R**esources
D : **D**ata Curation
O : Writing - **O**riginal Draft
E : Writing - Review & **E**diting
Vi : **Vi**sualization
Su : **Su**pervision
P : **P**roject administration
Fu : **Fu**nding acquisition

CONFLICT OF INTEREST STATEMENT

Authors state no conflict of interest.

ETHICAL APPROVAL

The research did not involve human participants or animals; therefore, ethical approval was not required for this study.

DATA AVAILABILITY

The D2-mec dataset analyzed during the current study is available from the study cited as [46]. The Csim dataset generated during this study is available from the corresponding author, Y.C. Wu, upon reasonable request.

REFERENCES

[1] P. Oltermann, "An employee shows shredded records of the Stasi," *The Guardian*, 2018. https://www.theguardian.com/world/2018/jan/03/stasi-files-east-germany-archivists-losing-hope-solving-worlds-biggest-puzzle (accessed Dec. 01, 2025).

[2] D. Heingartner, "Back together again," *The New York Times*, 2003. http://www.nytimes.com/2003/07/17/technology/back-together-again.html

[3] C. Schauer, M. Prandtstetter, and G. R. Raidl, "A memetic algorithm for reconstructing cross-cut shredded text documents," in *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 2010, vol. 6373 LNCS, pp. 103–117. doi: 10.1007/978-3-642-16054-7_8.

[4] N. Parvin and P. Kavitha, "Content based image retrieval using feature extraction in JPEG domain and genetic algorithm," *Indonesian Journal of Electrical Engineering and Computer Science*, vol. 7, no. 1, pp. 226–233, 2017, doi: 10.11591/ijeecs.v7.i1.pp226-233.

[5] W. Park and J. Ryu, "Fine-grained self-supervised learning with jigsaw puzzles for medical image classification," *Computers in Biology and Medicine*, vol. 174, 2024, doi: 10.1016/j.compbiomed.2024.108460.

[6] J. Joo, H. Lee, and S. C. Kim, "GMS-JIGNet: guided multi-scale jigsaw puzzles for self-supervised artificial spot segmentation in fundus photography," *Scientific Reports*, vol. 15, no. 1, 2025, doi: 10.1038/s41598-025-07077-4.
[7] M. M. Abdulrazzaq *et al.*, "Consequential advancements of self-supervised learning (SSL) in deep learning contexts," *Mathematics*, vol. 12, no. 5, p. 758, Mar. 2024, doi: 10.3390/math12050758.
[8] X. Song, J. Jin, C. Yao, S. Wang, J. Ren, and R. Bai, "Siamese-Discriminant Deep Reinforcement Learning for Solving Jigsaw Puzzles with Large Eroded Gaps," in *Proceedings of the 37th AAAI Conference on Artificial Intelligence, AAAI 2023*, 2023, vol. 37, pp. 2303–2311. doi: 10.1609/aaai.v37i2.25325.
[9] M. Huang, "Reconstruction of shredded figures based on grayscale matrix," *International Journal of Pattern Recognition and Artificial Intelligence*, 2025, doi: 10.1142/s0218001425540242.
[10] C. Ran, X. Li, and F. Yang, "Multi-step structure image inpainting model with attention mechanism," *Sensors*, vol. 23, no. 4, 2023, doi: 10.3390/s23042316.
[11] Y. Zhao and M. Su, "A puzzle solver and its application in speech descrambling," in *CEA'07 Proceedings of the 2007 annual Conference on International Conference on Computer Engineering and Applications*, 2007, no. d, pp. 171–176. [Online]. Available: http://dl.acm.org/citation.cfm?id=1348289
[12] T. Chuman and H. Kiya, "A jigsaw puzzle solver-based attack on image encryption using vision transformer for privacy-preserving DNNs," *Information (Switzerland)*, vol. 14, no. 6, 2023, doi: 10.3390/info14060311.
[13] P. Butler, P. Chakraborty, and N. Ramakrishan, "The Deshredder: A visual analytic approach to reconstructing shredded documents," in *IEEE Conference on Visual Analytics Science and Technology 2012, VAST 2012 - Proceedings*, 2012, pp. 113–122. doi: 10.1109/VAST.2012.6400560.
[14] A. Dong and V. Melnykov, "A stochastic optimization approach for shredded document reconstruction in forensic investigations," *Journal of Forensic Sciences*, vol. 71, no. 1, pp. 492–508, 2026, doi: 10.1111/1556-4029.70182.
[15] V. V. Nabiyev, S. Yılmaz, A. Günay, G. Muzaffer, and G. Ulutaş, "Shredded banknotes reconstruction using AKAZE points," *Forensic Science International*, vol. 278, pp. 280–295, Sep. 2017, doi: 10.1016/j.forsciint.2017.07.014.
[16] M. O. Omolara, M. A. Abah, M. B. Mohammed, and A. J. Ocheineh, "Anthropometry in forensic skeletal analysis: A critical review of its role in human identification," *Science International*, vol. 13, no. 1, pp. 110–120, 2025, doi: 10.17311/sciintl.2025.110.120.
[17] E. Pietroni and D. Ferdani, "Virtual restoration and virtual reconstruction in cultural heritage: Terminology, methodologies, visual representation techniques and cognitive models," *Information (Switzerland)*, vol. 12, no. 4, 2021, doi: 10.3390/info12040167.
[18] E. Iakovaki, D. Makris, and E. Malea, "Digital documentation and reconstruction of fragile organic artefacts: A state-of-the-art review," in *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences - ISPRS Archives*, 2025, vol. 48, no. M-9–2025, pp. 607–614. doi: 10.5194/isprs-archives-XLVIII-M-9-2025-607-2025.
[19] C. Ostertag and M. Beurton-Aimar, "Using graph neural networks to reconstruct ancient documents," in *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 2021, vol. 12667 LNCS, pp. 39–53. doi: 10.1007/978-3-030-68787-8_3.
[20] R. Comes, C. G. D. Neamțu, C. Grec, Z. L. Buna, C. Găzdac, and L. Mateescu-Suciu, "Digital reconstruction of fragmented cultural heritage assets: The case study of the Dacian embossed disk from Piatra Roșie," *Applied Sciences (Switzerland)*, vol. 12, no. 16, 2022, doi: 10.3390/app12168131.
[21] X. Y. Zhang, J. H. Yang, Y. J. Gong, Z. H. Zhan, and J. Zhang, "Tree structured cooperative coevolutionary genetic algorithm for fragment reconstruction," *IEEE Transactions on Evolutionary Computation*, 2025, doi: 10.1109/TEVC.2025.3550742.
[22] L. Zhang, Z. Wang, Y. Dong, and Y. Zhu, "CG2AN: Conditional graph generative adversarial networks for shredded image restoration," in *Frontiers in Artificial Intelligence and Applications*, 2025, vol. 413, pp. 1985–1992. doi: 10.3233/FAIA251034.
[23] F. Yan, Y. Zheng, J. Cong, L. Liu, D. Tao, and S. Hou, "Solving jigsaw puzzles via nonconvex quadratic programming with the projected power method," *IEEE Transactions on Multimedia*, vol. 23, pp. 2310–2320, 2021, doi: 10.1109/TMM.2020.3009501.
[24] W. Kong and B. B. Kimia, "On solving 2D and 3D puzzles using curve matching," in *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 2001, vol. 2, pp. 583–590. doi: 10.1109/cvpr.2001.991015.
[25] E. Justino, L. S. Oliveira, and C. Freitas, "Reconstructing shredded documents through feature matching," *Forensic Science International*, vol. 160, no. 2–3, pp. 140–147, 2006, doi: 10.1016/j.forsciint.2005.09.001.
[26] C. Solana, E. Justino, L. S. Oliveira, and F. Bortolozzi, "Document reconstruction based on feature matching," in *Brazilian Symposium of Computer Graphic and Image Processing*, 2005, vol. 2005, pp. 163–170. doi: 10.1109/SIBGRAPI.2005.26.
[27] M. G. Chung, M. M. Fleck, and D. A. Forsyth, "Jigsaw puzzle solver using shape and color," in *International Conference on Signal Processing Proceedings, ICSP*, 1998, vol. 2, pp. 877–880. doi: 10.1109/icosp.1998.770751.
[28] F. H. Yao and G. F. Shao, "A shape and image merging technique to solve jigsaw puzzles," *Pattern Recognition Letters*, vol. 24, no. 12, pp. 1819–1835, 2003, doi: 10.1016/S0167-8655(03)00006-0.
[29] P. De Smet, J. De Bock, and W. Philips, "Semiautomatic reconstruction of strip-shredded documents," in *Image and Video Communications and Processing 2005*, 2005, vol. 5685, p. 239. doi: 10.1117/12.586340.
[30] M. Alex Oliveira Marques and C. Obladen Almendra Freitas, "Document decipherment-restoration: Strip-shredded document reconstruction based on color," *IEEE Latin America Transactions*, vol. 11, no. 6, pp. 1359–1365, 2013, doi: 10.1109/TLA.2013.6710384.
[31] G. Chen, J. Wu, C. Jia, and Z. Yunzhou, "A pipeline for reconstructing cross-shredded English document," in *2017 2nd International Conference on Image, Vision and Computing, ICIVC 2017*, 2017, pp. 1034–1039. doi: 10.1109/ICIVC.2017.7984711.
[32] H. Y. Lin and W. C. Fan-Chiang, "Reconstruction of shredded document based on image feature matching," *Expert Systems with Applications*, vol. 39, no. 3, pp. 3324–3332, 2012, doi: 10.1016/j.eswa.2011.09.019.
[33] M. Prandtstetter and G. R. Raidl, "Combining forces to reconstruct strip shredded text documents," in *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 2008, vol. 5296 LNCS, pp. 175–189. doi: 10.1007/978-3-540-88439-2_13.
[34] J. Balme, "Reconstruction of shredded documents in the absence of shape information," USA, 2007.
[35] W. Morandell, "Evaluation and reconstruction of strip-shredded text documents," Institute of Computer Graphics and Algorithms, Vienna University of Technology, 2008. [Online]. Available: https://www.ads.tuwien.ac.at/publications/bib/pdf/morandell_08.pdf
[36] P. Li, X. Fang, L. Pan, Y. Piao, and M. Jiao, "Reconstruction of shredded paper documents by feature matching," *Mathematical Problems in Engineering*, vol. 2014, pp. 1–9, 2014, doi: 10.1155/2014/514748.
[37] J. Perl, M. Diem, F. Kleber, and R. Sablatnig, "Strip shredded document reconstruction using optical character recognition," in *4th International Conference on Imaging for Crime Detection and Prevention 2011, ICDP 2011*, 2011, pp. 35–41. doi: 10.1049/ic.2011.0132.

[38] T. M. Paixao, M. C. S. Boeres, C. O. A. Freitas, and T. Oliveira-Santos, "Exploring character shapes for unsupervised reconstruction of strip-shredded text documents," *IEEE Transactions on Information Forensics and Security*, vol. 14, no. 7, pp. 1744–1754, 2019, doi: 10.1109/TIFS.2018.2885253.
[39] D. Sholomon, O. E. David, and N. S. Netanyahu, "Dnn-buddies: A deep neural network-based estimation metric for the jigsaw puzzle problem," in *Lecture Notes in Computer Science*, 2016, vol. 9887 LNCS, pp. 170–178. doi: 10.1007/978-3-319-44781-0_21.
[40] T. M. Paixão *et al.*, "Self-supervised deep reconstruction of mixed strip-shredded text documents," *Pattern Recognition*, vol. 107, p. 107535, 2020, doi: 10.1016/j.patcog.2020.107535.
[41] J. Sauvola and M. Pietikäinen, "Adaptive document image binarization," *Pattern Recognition*, vol. 33, no. 2, pp. 225–236, 2000, doi: 10.1016/S0031-3203(99)00055-2.
[42] F. N. Iandola, S. Han, M. W. Moskewicz, K. Ashraf, W. J. Dally, and K. Keutzer, "SqueezeNet: AlexNet-level accuracy with 50x fewer parameters and <0.5MB model size," *arXiv:1602.07360*, Nov. 2016.
[43] D. P. Kingma and J. L. Ba, "Adam: A method for stochastic optimization," in *3rd International Conference on Learning Representations, ICLR 2015 - Conference Track Proceedings*, 2015, pp. 1–15.
[44] V. Chvátal, D. Applegate, R. Bixby, and W. Cook, "CONCORDE: A code for solving traveling salesman problems." 1999. [Online]. Available: http://www.tsp.gatech.edu/concorde.html
[45] R. Jonker and T. Volgenant, "Transforming asymmetric into symmetric traveling salesman problems: erratum," *Operations Research Letters*, vol. 5, no. 4, pp. 215–216, 1986, doi: 10.1016/0167-6377(86)90081-7.
[46] T. M. Paixão, R. F. Berriel, M. C. S. Boeres, C. Badue, A. F. De Souza, and T. Oliveira-Santos, "A deep learning-based compatibility score for reconstruction of strip-shredded text documents," in *Proceedings - 31st Conference on Graphics, Patterns and Images, SIBGRAPI 2018*, 2018, pp. 87–94. doi: 10.1109/SIBGRAPI.2018.00018.
[47] T. A. Nartker, S. V. Rice, and S. E. Lumos, "Software tools and test data for research and testing of page-reading OCR systems," in *Document Recognition and Retrieval XII*, Jan. 2005, vol. 5676, pp. 37–47. doi: 10.1117/12.587293.
[48] C. Le, N. Deng, and X. Jin, "Reconstructing regular shredded documents by similarity measure," in *Proceedings of the 2nd International Conference on Soft Computing in Information Communication Technology*, 2014, vol. 71, pp. 160–162. doi: 10.2991/scict-14.2014.38.
[49] P. De Smet, "Semi-automatic forensic reconstruction of ripped-up documents," in *Proceedings of the International Conference on Document Analysis and Recognition, ICDAR*, 2009, pp. 703–707. doi: 10.1109/ICDAR.2009.7.
[50] F. Alhaj, A. Sharieh, and A. Sleit, "Reconstructing colored strip-shredded documents based on the Hungarians algorithm," in *2019 2nd International Conference on New Trends in Computing Sciences, ICTCS 2019 - Proceedings*, 2019, pp. 1–6. doi: 10.1109/ICTCS.2019.8923048.

BIOGRAPHIES OF AUTHORS

Yi-Chang Wu works for the Ministry of Justice Investigation Bureau. He received his M.E.E degree in communications engineering from the National Sun Yat-sen University, Taiwan, in 2007, and the Ph.D. in electronic and computer engineering from National Taiwan University of Science and Technology, Taiwan, in 2019. His research interests cover various aspects of machine learning, robotics, image analysis and monitoring laboratory. He can be contacted at address is shintenwu@gmail.com.

Pei-Shang Chiang works for the Ministry of Justice Investigation Bureau. She received her M.C.E degree in Earth Science from National Cheng Kung University, Taiwan, in 2003. Her current research interests are machine vision and image analysis. She can be reached via email address: carriechiang0414@gmail.com.

Yao-Cheng Liu was born in Kaohsiung, Taiwan. He is pursuing an M.E.E degree in computer and communication at Jinwen University of Science and Technology. His research interests are AI literacy, emotion and anthropomorphism. He is working for the Ministry of Justice Investigation Bureau in Taiwan. The email address is 40981t1371@gmail.com.