Manuscripts fidelity in the digital libraries era: the contrast enhancement evaluation conundrum

Martina Franchi^{1,2,†}, Tiziana Cattai^{3,†}, Stefania Colonnese^{3,*,†} and Azeddine Beghdadi^{4,†}

¹Department of Basic and Applied Sciences for Engineering, Sapienza University of Rome, Italy

²CNR-Institute of Nanotechnology,c/o Physics Department, Sapienza University of Rome, Italy

³Department of Information Engineering, Electronics and Telecommunications, Sapienza University of Rome, Italy

⁴Institut Galilée, Sorbonne Paris Nord, Paris, France

Abstract

Digitization of ancient manuscripts spurs inter-disciplinary research but the quality may be limited due to manuscript preservation state or scan hardware limitations. Contrast Enhancement improves the subjective quality of images for end users. Estimating the performance of contrast enhancement is challenging since ancient manuscript are composite, i.e. they contain drawings and calligraphic elements, and the contrast enhancement metric should capture fidelity, color quality and recovery of faded text. After applying different global and local contrast enhancement techniques to a set of 15th century manuscripts, where the text and pictorial representations were partially compromised due to conservation problems, we assessed the quality of the enhanced manuscripts by performance metrics, and compared them with human supervised ranking of the enhanced manuscript in terms of either color and text quality. By comparison of the metrics with the supervised ranking results, we identify the most accurate performance metric, namely the metric based on brightness preservation. Future work will address evaluation of the enhancement for artificial intelligence based segmentation, dating, visual search, text recognition purposes.

Keywords

contrast, manuscript, quality metric, readability

1. Introduction

Digital acquisition of ancient manuscripts makes them available by virtual libraries around the world and spurs inter-disciplinary research [1, 2]. Digitized manuscript subjective quality [3] may be limited due to the manuscript state of preservation, to the limitations of the camera or scan used in the acquisition procedure, or due to image compression. Therefore, enhancement is applied [4] to improve the subjective quality of images for end users. Enhancing the digitized manuscript contrast, i.e. the psycho-physical effect of the visual stimulus [5, 6, 7, 8], changes the image feature possibly affecting its perceived quality [9].

For natural or composite images, the perceived quality [10] of virtual images is predicted with image quality assessment techniques, with different focuses such as automatic assessment [11], multi-camera setup [12], perceptual features [13], image layout [14], multimodal data [15]. For enhanced images, suitable metrics shall be defined [16].

Focusing on Contrast Enhancement (CE) of ancient manuscripts, we aim to identify the most suited state-of-the-art metric for CE Evaluation (CEE). Estimating the performance of nonlinear processing techniques as contrast enhancement is an open research topic [17]. Besides, ancient manuscripts are composite, i.e. they contain drawings and calligraphic elements. Therefore, the CEE metric should capture CE performances in terms of color quality and recovery of the faded text. Our goal is therefore

- [†]These authors contributed equally.

D 0000-0002-1807-2155 (.S. Colonnese)

CELIB

CVCS2024: the 12th Colour and Visual Computing Symposium, September 5–6, 2024, Gjøvik, Norway *Corresponding author.

[🛆] martina.franchi@uniroma1.it (M. Franchi); tiziana.cattai@uniroma1.it (T. Cattai); stefania.colonnese@uniroma1.it (.S. Colonnese); azeddine.beghdadi@univ-paris13.fr (A. Beghdadi)

^{© 02024} Copyright for this paper by its authors. Use permitted under Creative Commons License Attribution 4.0 International (CC BY 4.0).

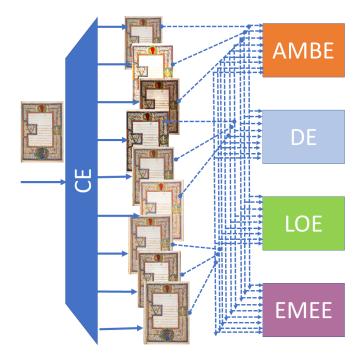


Figure 1: Overall view of the approach: different Contrast Enhancement (CE) methods are applied to the original digitized manuscript; the enhanced copies are evaluated using different state of the art CE Evaluation (CEE) metrics. The CEE results are compared with supervised manuscript ranking.

twofold, since we assess both the visual quality of the enhanced manuscript and the readability of its content.

We analyze a set ancient manuscripts from 15th century, that had conservation issues and in which the text and pictorial representations were partly compromised. After applying different global and local contrast enhancement techniques, we i) assess the quality of the enhanced manuscripts by CEE metrics, and ii) perform two supervised rankings of the enhanced manuscript, addressing both the color quality and text readability. Then, we compare the results of the CEE metrics with the supervised CE ranking and identify the metric that more consistently matches the supervised CE ranking. In these results we show that the highest ranked method is the one that mostly preserves brightness in the enhancement [18]. For the sake of completeness, after assessing how contrast enhancement changes the subjective quality of the text, we also provide an example to discuss how it affects the readability by an OCR algorithm, since automatic text extraction would boost information mining in digitized libraries ancient manuscript.

We present in Fig. 1 the graphical abstract of our proposed approach.

2. Related work

Manuscripts are susceptible to deterioration, particularly when not adequately preserved over time and exposed to atmospheric or anthropic agents, which accelerate natural degradation. Several studies focus on analyzing the materials used in manuscripts (such as ink, dyes, and medium), often employing multi-analytical spectroscopic techniques [24], [25]. In the context of cataloging in virtual libraries, digitization facilitates processes like blind source separation, self-organizing maps extraction, and linear discriminant analysis, led to reveal hidden features in the manuscripts [25].

In the field of virtual restoration, recent literature has introduced innovative methods [26], such as color-based segmentation combined with Gaussian Mixture Models, aimed at enhancing readability. Additionally, in digital text restoration, researchers[3] have presented methods to discern human preferences for legible text using datasets containing texts from various manuscripts. Building upon these approaches, our study initially focuses on restoring colors and text in digital RGB images of



Figure 2: Effect of CE methods on the original manuscript [19]: (first row) original, Amplitude Scaling (AS) [20], Histogram Equalization (HE) [11], Contrast Limited Adaptive HE (CLAHE) [12] and Linear Unsharp Masking (UM) [21]; (second row) Local Laplacian filtering (LLF)[22], Fast Low-lighting Enhancement (FLE) [23], FLE with Dark Priors (FLE-DP) [9], Local Laplacian filtering LLF (heavy, $\sigma = 0.2$, $\alpha = 0.3$) [13] and UM with Gaussian LP filtering (UMLP)[21].

manuscripts using contrast enhancement techniques in the spatial domain. Enhancement methods can be applied also to UV images and infrared images [27], giving good results in specific cases to recover faded text.

In the current state of the art, contrast enhancement techniques on digital images vary based on emphasized image features and the local or global approach used. There are also learning-based methods that do not identify features to emphasize [28]. The evaluation of contrast enhancement by metrics can employ subjective or objective approaches. There are also metrics based on machine learning [29]. Objective approaches provide metrics that can be applied using algorithms and are different: statistic-based, gradient-energy-based, and human vision system (HVS) inspired. In our work, we focus on picking one metric for both statistic-based and HVS-inspired approaches.

The contrast enhancement metrics can be categorized as Full-Reference Image Quality Assessment (FR-IQA) or Reduced- and No-Reference Image Quality Assessment [30]. We use both approaches, Full-Reference and No-Reference. Digitized manuscripts are a particular class of image that can present text and composed figures, and their wide variability in manuscript layout and textual formats pose challenges for deep learning networks. Additionally, texts in different languages and partially faded portions in the dataset complicate the training process.

Furthermore, reproducing data augmentation to simulate degradation and distortion in the image, including layout and text, can also be challenging. The enhancement for the text present in the image can also be assessed by OCR (optical character recognition) that uses deep learning. There are several studies about the transcription of manuscripts in different languages [31]. In the contemporary research landscape, ongoing discussions persist regarding the methodologies for contrast enhancement and its subsequent evaluation. Our study represents an application tailored for a specific dataset within a virtual library setting. This dataset poses unique challenges, rendering it particularly resistant to analysis using machine learning methodologies. However, despite these challenges, the potential of machine learning-based approaches can be probed in the performance of contrast enhancement and its subsequent evaluation.

3. Assessing contrast enhancement of ancient manuscripts

The main goal of this paper is to compare different CEE metrics in terms of their ability to assess the quality of enhanced ancient manuscripts regarding both color and readability, and to compare these assessments with those conducted through subjective evaluation. In this section, we introduce state-of-the-art CE methods to be applied to our dataset, the selected CEE metrics aimed at investigating their performance on enhanced manuscript images, and the methodology employed for a subjective contrast evaluation of these images.

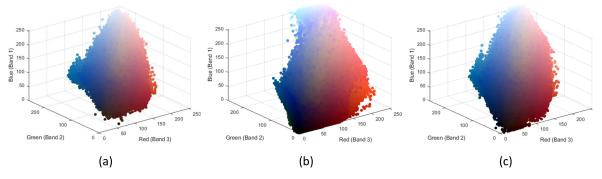


Figure 3: Scatter plots illustrating the RGB color space of an identical input image: (a) original, (b) enhanced using the HE method, and (c) enhanced using UM, with colors representing RGB values [19].

3.1. Contrast enhancement methods

State-of-the-art CE methods either directly adopt a contrast definition or indirectly alter the contrast by modifying a few image features. Besides, they leverage either global characteristics of the images (i.e. average of maximum and minimum luminance values of all small fractions of the image) [32], or local contrast measures, based on characteristics measured on neighboring pixels [33].

Among several state-of-the-art algorithms, we selected nine CE methods, reported in Table 1 and fixed their parameters.

Table 1

CE Methods

1	Amplitude Scaling (AS)	[20]
2	Histogram Equalization (HE)	[11]
3	Contrast Limited Adaptive HE (CLAHE)	[12]
4	Linear Unsharp Masking (UM)	[21]
5	Local Laplacian filtering (LLF, $\sigma = 0.4, \alpha = 0.5$)	[22]
6	Fast Low-lighting Enhancement (FLE)	[23]
7	FLE with Dark Priors (FLE-DP)	[9]
8	Local Laplacian filtering (LLF, $\sigma = 0.2, \alpha = 0.3$)	[13]
9	UM with Gaussian LP filtering (UMLP, $R = 1.2$)	[21]

We illustrate the effect of the Contrast Enhancement methods on one of the ancient manuscripts shown in Fig. 2, namely a manuscript containing a poem in honor of Sultan Mehmed II by Gian Mario FileIfo from collections of Bibliothèque de Genève. It is characterized by conservation issues that result in loss of colour in the region of the decorated frame and the first lines of the text. The CE methods result in very different enhanced manuscripts.

Fig.3 shows how the CE algorithms act on the manuscript color space. In Fig.3(a) each pixel of the original digitized manuscript is represented as a point of the RGB space. Fig.3(b) and (c) represent

the pixels of the manuscript enhanced by HE and UM, respectively. The distribution of the pixels in the RGB space is significantly altered by HE, that tends to occupy a larger RGB volume, whereas UM basically maintains the original pixels distribution with a slight amplification of the color components.

3.2. Contrast enhancement evaluation metrics

A variety of contrast enhancement metrics have been defined in the literature [16, 9]. We chose classical Contrast Enhancement Evaluation (CEE) metrics to apply to different enhanced manuscript to identify the metric that best captures the enhancement performance.

We have selected four metrics for Contrast Enhancement Evaluation. These metrics focus on various aspects of the image that may be affected after contrast enhancement (CE), with each being sensitive to different artifacts introduced by CE. Based on this, CEE metrics can be divided into two groups: Full Reference (FR) and No Reference (NR) measures. If the original image is taken into account and a comparison between pre- and post-enhancement is made, we have full reference measures; if only the post-enhanced image is evaluated, we have no reference measures.

Furthermore, these metrics can be classified into different classes based on the approach used to assess the image for CE evaluation, such as Statistics-based, Human Visual System (HVS)-inspired, and Gradient/Energy-based. In our case, we selected the first two classes. For the Statistics-based approach, we have chosen the Absolute Mean Brightness Error (AMBE), Discrete Entropy (DE), and Lightness Order Error (LOE) metrics. For the Human Visual System (HVS)-inspired approach, we selected the Measure of Enhancement by Entropy (EMEE) metric. Contrast assessment is usually performed on the luminance component.

Absolute Mean Brightness Error (AMBE) (FR) [18] measures the difference between the mean value of the L^{*} component of the enhanced image, i.e. the brightness degradation in the enhanced image.

$$AMBE = \left| \mathbb{E}_{ij} \{ I_{ij}^{(or)} \} - \mathbb{E} \{ I_{ij}^{(enh)} \} \right|$$
(1)

Lightness Order Error (LOE)(NR) [16] measures the changes of the lightness \tilde{I} between the original and enhanced manuscript, being the lightness defined as the pixel-wise maximum of the three color image components, namely $\tilde{I}_{ij} = \max(R_{ij}, G_{ij}, B_{ij})$.

$$LOE = \mathbb{E}_{ij} \left\{ \sum_{uv} \left(\tilde{I}_{ij}^{(or)} > \tilde{I}_{uv}^{(or)} \right) \oplus \left(\tilde{I}_{ij}^{(enh)} > \tilde{I}_{uv}^{(enh)} \right) \right\}$$
(2)

Discrete Entropy (DE) (NR) [34] measures of the grade of randomness of gray-level of the enhanced image, assuming that CE leads to more visible details, implying the use of more gray levels.

$$DE = -\sum_{l} p(I_{ij}^{(enh)} = l) \log_2(p(I_{ij}^{(enh)} = l))$$
(3)

Measure of Enhancement by Entropy (EMEE) (NR) [6] measures the entropy of the enhanced image L^* component as an indicator of the spatial content information.

$$EMEE = \mathbb{E}_{ij} \left\{ \frac{\max\left(I_{ij}^{(enh)}\right)}{\min\left(I_{ij}^{(enh)}\right) + \epsilon} \cdot \log\left(\frac{\max\left(I_{ij}^{(enh)}\right)}{\min\left(I_{ij}^{(enh)}\right) + \epsilon}\right) \right\}$$
(4)

We compute the AMBE, LOE, DE, and EMEE metrics and we map them into quality scores Q_{AMBE} , Q_{DE} , Q_{LOE} , and Q_{EMEE} defined as follows:

$$Q_{AMBE} = 1 - \frac{AMBE}{M_{AMBE}}; \ Q_{LOE} = 1 - \frac{LOE}{M_{LOE}};$$

$$Q_{DE} = \frac{DE}{M_{DE}}; \ Q_{EMEE} = \frac{EMEE}{M_{EMEE}}$$
(5)

being M_{AMBE} , M_{LOE} , M_{DE} , M_{EMEE} the maximum values assumed by the metrics on the 9 enhanced manuscripts. All the quality scores Q_{AMBE} , Q_{DE} , Q_{LOE} , and Q_{EMEE} range in [0, 1], with 1 corresponding to the maximum quality.

3.3. Subjective evaluation of contrast enhancement

The contrast enhancement evaluation is also assessed using subjective methods based on human judgment of the perceived quality of an image. Subjective methods are more reliable for judging the quality of an image, as the ultimate goal is for the image to be visually appreciated by humans. Vision is a complex process, and CEE metrics may not accurately reflect human judgment if they are not based on the human visual system (HVS). The correspondence observed between the subjective ranking and the objective scores will provide insight into the degree of consistency between a given CEE metric and human judgment regarding the quality of the images. The test procedures to be followed have been based on the recommendations set out in the [35], [36]. In a subjective experiment, at least 15 different subjects assess the quality of images based on their perception. Subjective experiment methods can be classified into two groups: rating-based and ranking-based methods [9]. In rating-based methods, each subject assigns a score to each stimulus on an interval scale or categorical scale. There are several rating-based methods, which can be classified into three categories: Single-Stimulus (SS), Double-Stimulus (DS), and Multi-Stimulus methods. We will not delve into the explanation of these methods in this paper. The ranking-based methods can be divided into two groups: rank order-based methods and Pairwise Comparison (PC)-based methods [36]. In rank order-based methods, the subjects rank different stimuli displayed at once according to the perceived quality. In PC-based methods, the subjects observe stimuli presented in pairs and choose which one they prefer or if they are alike. For this study, we conducted a subjective assessment of contrast enhancement using a ranking-based method, specifically the Pairwise Comparison (PC)-based method. We provided instructions to the observers to evaluate the quality of the enhanced versions of 12 medieval manuscripts. For each manuscript, they were asked to open the links containing images of the manuscript and indicate the best version (rated from 0 to 9) for readability and the best version (also rated from 0 to 9) for pleasant chromatism. We organized 6 groups, each consisting of 1 to 4 persons, and tasked them with identifying the best out of 9 enhanced manuscripts in terms of readability and color quality for each of the 12 original manuscripts. To simulate a realistic scenario, some groups visualized the enhanced manuscripts on a personal computer, while others used a mobile device. Subsequently, we converted the ranking results into a subjective score for each contrast enhancement method, defined as the count of how many times the corresponding enhanced manuscript was selected as the best.

4. Experimental results

Herein, we firstly analyze the CEE metrics and compare them with the supervised CE rankings collected over the set of manuscripts. Then, we are able to identify the metric that more consistently matches the supervised CE ranking. Finally, for the sake of completeness, we present a toy case showing the contrast enhancement impact on automatic text recognition.

4.1. Contrast enhancement metric and manuscript subjective quality

We consider a set of 12 digitized manuscripts, illustrated in Fig.5, from different collections and available through the virtual library [37]. The digitization hardware includes Hasselblad model H4D-50MS (50 Megapixel, Multi-Shot) cameras and the Hasselblad medium format camera, Model H3DII-31 (31 Megapixel). The digitized manuscript are in the JPEG 2000 format, which is adopted for both the web application viewer and the image server. We downloaded JPEG "small" size from a five-choice due to computational time for images processing. These manuscripts present certain conservation issues, such as the loss of color in the text, which makes them difficult to read. This problem also affects the clarity of the pictures and the border decorations. As shown in the pipeline in Figure 1, we processed

manuscript images using 9 CE methods. Subsequently, the enhanced images underwent evaluation using four Contrast Enhancement Evaluation metrics (CEE). The results of the CEE metrics were then normalized using the equations (5) to facilitate comparison among them.

Fig.4 shows the correlation of the CEE values assumed by the 4 metrics over the 9 CE methods for the manuscript [19]. As a first result, we observe that the metrics are scarcely correlated, i.e. they do not provide coherent results in scoring the CE methods.

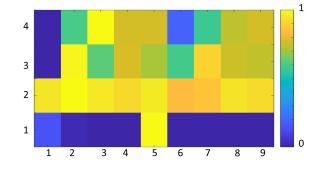


Figure 4: Correlation among the 4 metrics score over the 9 methods for the manuscript [19].

We then extend the analysis to the set of manuscript images in Fig.5. Fig.6 shows the polar plots of the 4 normalized evaluation metrics versus the 9 enhancement methods. Specifically, Fig.6(a)-(d) represents Q_{AMBE} , Q_{DE} , Q_{LOE} , and Q_{EMEE} , respectively. We recognize that Q_{LOE} presents a small sensitivity to the CE differences, whereas Q_{EMEE} have drastic oscillations. In order to identify the metric best matching the subjective assessment, a supervised ranking of the enhanced manuscript has been realized. The ranking has been obtained by human subjective visual analysis, as mentioned in the section 3.3, We then converted their rankings into subjective scores based on the frequency of selection for each enhanced manuscript image. In Fig.7 we compare the score computed by the supervised subjective ranking results with the Q_{AMBE} , Q_{DE} , Q_{LOE} , and Q_{EMEE} metrics. The bar plot in Fig.7(a) summarizes the average score, computed over the 12 manuscripts, of the 4 CE quality metrics Q_{AMBE} , Q_{DE} , Q_{LOE} for the 9 considered enhanced methods; the standard deviation of the metrics is also shown by a red vertical line. The bar plot in Fig.7(b) reports a subjective score for each CE method, both in terms of readability and color quality. We recognize that the method of collecting subjective choices results in a consistent calculation of the subjective score, as it is easier for the user to select the best looking improved manuscript than to assign an individual score to each manuscript. Figure 7(b) shows that the most effective method for subjective evaluation is method 3, which corresponds to the CLAHE method. Referring to Fig. 6, which displays the results of the 4 CEE metrics, the colored arrows point to enhanced method number 3. This method obtained the highest ratings in subjective evaluations for readability (indicated by the purple arrow) and color quality (indicated by the green arrow) among the processed images. It is observed that Q_{AMBE} aligns with the subjective evaluation, demonstrating a good normalized score of 1 for the CLAHE metric. Additionally, the Q_{LOE} method records a good ranking, but it shows to be a metric less sensitive to various conditions and image artifacts.

In conclusion, from these results it stems that in ancient manuscript contrast enhancement the preservation of the brightness, as measured by the normalized CEE Q_{AMBE} , may be a relevant factor affecting the manuscript quality ranking. This indicates that Q_{AMBE} correlates closely with subjective ranking and demonstrates sensitivity to variations in image content. It's clear that different image content leads to diverse enhancement outcomes. Therefore, metrics that are sensitive to differences in layout and artifacts introduced are crucial for effectively assessing quality. It's worth noting that exploring a broader range of CEE metrics could be beneficial in future research. Additionally, the presented methodology lays the groundwork for extending to various supervised ranking methods, considering the suitability of enhanced manuscripts for AI processing tasks like image or text retrieval, visual search, and dating.



Figure 5: Original manuscripts from collections of Bibliothèque de Genève, (https://www.e-codices.ch), [38] (754 × 539), [39] (907 × 696), [40] (959 × 627), [41] (819 × 642), [42] (570 × 392), [43] (578 × 366), [44] (570 × 392), [45] (768 × 518), [46] (821 × 660), [47] (745 × 560), [19] (937 × 659), [48] (756 × 537)

4.2. Contrast enhancement metric and automatic text readability

To estimate how much the applied contrast methods improve character visibility and, therefore, readability, we consider Optical Character Recognition (OCR). Recently, OCR for digitizing handwritten documents are improved and they are principally based on deep learning [31]. While many datasets in different modern languages are built to train the OCR, OCR studies for Latin language are still limited. Additionally, for proper recognition, OCR need to be trained to recognize characters under various conditions such as background illumination, camera angle, distortion [49], and curved text [50],[51]. These aspects severely affect current OCR performances, preventing direct application of on-the-shelf OCR algorithms to actual manuscripts. Therefore, we test the CE methods on the toy case or English handwritten text, representing the most studied language. The original handwritten text is shown in Fig. 8(a). We altered it to resemble the images of treated manuscripts by i)changing the color of both

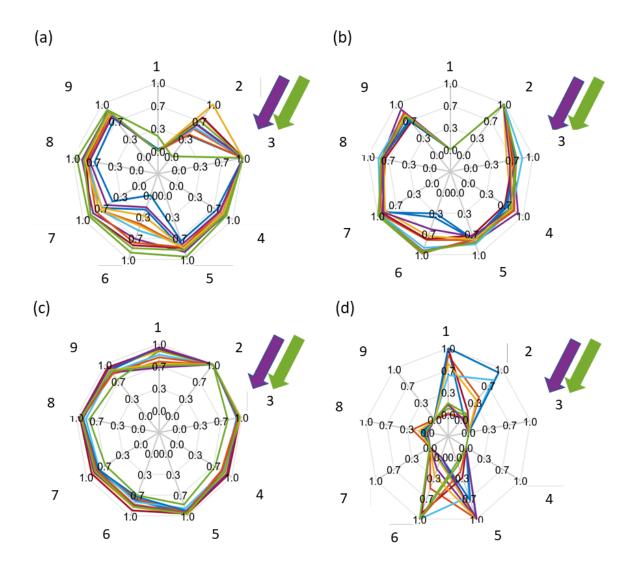


Figure 6: Polar plots of the normalized evaluation metrics ((a) Q_{AMBE} ,(b) Q_{DE} ,(c) Q_{LOE} ,(d) Q_{EMEE}) versus the 9 enhancement methods listed in Table 1. The arrows highlight the most voted method (number 3 (CLAHE)) from subjective evaluation in terms of readability (in purple) and color pleasure (in green). The subjective results are reported across all metrics. Each coloured line corresponds to a different manuscript of the set in Fig.5 from collections of Bibliothèque de Genève, (https://www.e-codices.ch), [38], [39], [40], [41], [42], [43], [44], [45], [46], [47], [19], [48].

background and text to brown tones, ii) introduced spatially varying blur across the entire image, and iii) adding salt and pepper noise, The degraded image is shown in Fig. 8(b). We then selected an online deep learning based OCR platform [52] specifically designed for handwritten text to assess the text readability. Firstly, we applied the OCR to the original and degraded text images: the OCR was able to extract 100% of the text from the clean image in Fig.8(a), whereas for the blur and noisy image in Fig. 8(b), the OCR only extracted the characters 'er'. Then, we applied the contrast enhancement methods to the blurred and noisy image in Fig. 8(b) obtaining nine enhanced images. Table 2 shows the transcript obtained for each contrast method applied. The AS method, with the parameter set in the previous study, does not produce any word transcriptions via OCR. We recognize that CLAHE[12], LLF σ = 0.4,

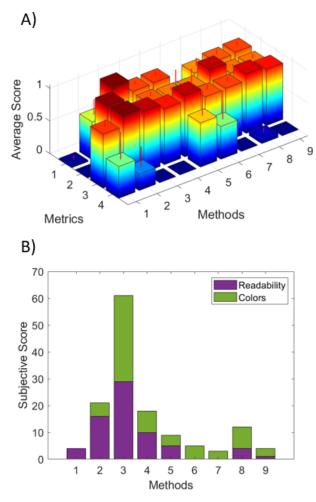


Figure 7: A) Average score, over the 12 manuscripts [38]-[48], of the 4 CE quality metrics Q_{AMBE} , Q_{DE} , Q_{LOE} for the 9 CE methods (the standard deviation is shown by a red vertical line); B subjective score for the 9 CE method (readability and color quality).

 $\alpha = 0.5[22]$, and LLF $\sigma = 0.2$, $\alpha = 0.3[13]$ are the most effective methods in improving readability. The LLF ($\sigma = 0.2$, $\alpha = 0.3$) method led to the enhanced image shown in Fig. 8(c), whose transcript ranked best with 28 words in common with the original text, corresponding to a similarity percentage of 58.33% between the two texts. A few remarks are in order. The OCR results show that the best contrast enhancement method for automatic character recognition (LLF) differ with respected to the one obtained from subjective rankings and CE metrics (CLAHE). Besides, on one hand the performance of the enhancement methods are highly dependent on the content of the image; on the other hand, the automatic character recognition is heavily affected by the training stage. To sum up, the processing of manuscript images pose unprecedented challenges to the development of CE algorithms and CEE metrics; future work is needed to develop spatially adaptive CEE metrics, accounting for both the readability of the text and visual quality of the figures.

5. Conclusion and future work

In this paper, we applied Contrast Enhancement (CE) techniques to ancient manuscript images sourced from a digital library [37]. These images contained text and representations affected by faded colors in certain areas. Among various traditional contrast enhancement methods, we selected nine that effectively improved salient characteristics (such as brightness and sharpness), aiming to enhance both the representation and text present in the manuscripts image. Our focus was to evaluate these enhancement methods using objective evaluations by Contrast Enhancement Evaluation (CEE) metrics, selecting

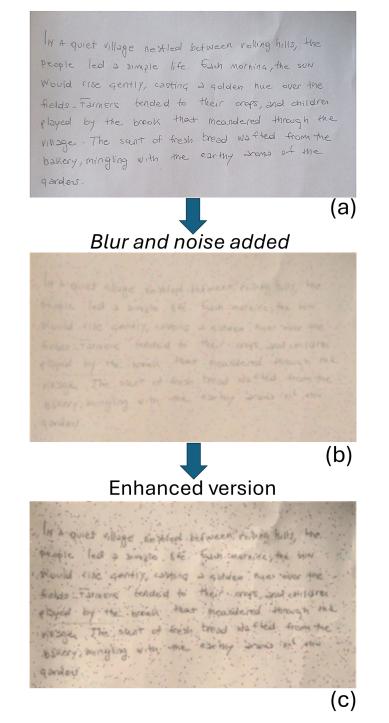


Figure 8: Test image for OCR: a. Image of handwritten English text; b. Image of handwritten text degraded with blur and noise; c. Enhanced image with heavy LLF filter.

among them both static-based and human visual system-inspired approach. In selecting the CE methods, our priority was to efficiently enhance the images, opting for methods that ensure both time efficiency and the preservation of color and structural information. Additionally, we complemented the objective evaluations with subjective rankings and automatic text recognition performance assessment. Our results revealed that different CEE metrics yielded varying outcomes depending on the image content, with limited correlation observed between. Moreover, the choice of contrast enhancement method significantly influenced CEE results, indicating sensitivity to applied CE techniques and associated artifacts produced. Importantly, we found a positive correlation between subjective preferences and

Table 2OCR transcription

1	(AS)	"_"
2	(HE)	
2	(пс)	"village de Hep between his 7
		gently, ating seyever the feld
		Farmers ended to crops, and are
		Measidered through The sont
		bread afted from from the overy,
		bungling exy gardey"
3	(CLAHE)	"quiet village and between rid-
		ing hulls, the ould fire gently,
		caring Sotage Aue over the elds
		Farm bonded to their crops, and
		played by oak that measidered
		through the The st read wafted
		from the sery, bungling qandey"
4	(UM)	bended their 4 bread to fled
		from the"
5	(LLF, $\sigma = 0.4, \alpha = 0.5$)	"quiet village settled between
		riding hulls, th people led a
		simple Suchmornine; the tow
		Rould rise gently, caping god
		yes over the fields Farmers
		tended to their crops, and
		I played by the break that
		neasidered through" The sunt
		fresh bread who fled from the
		musige bskery, bungling earthy
		brows"
6	(FLE)	tended to thei PHA bread do
L		fled from"
7	(FLE-DP)	"p Th"
8	(LLF, $\sigma = 0.2, \alpha = 0.3$)	"quiet village de test befween
		riding halls, the people led a
		simple life. Such mornite, the
		w fould rise gently, caving a
		golden nye over the fields Farm-
		ers tended to their crops, and
		played by the week that mean-
		dered through" The sunt of fresh
		bread who fled from the •wasge
		bakery, bringling with. earthy
		brows qanday"
9)	(UMLP, R = 1.2)	"Village between led a "tended
		to their crops, ered through
		bread wa to fled from the The"
L	1	1

CEE metrics, with the CLAHE method consistently ranked highly. In summary, Supervised subjective rankings, based on readability and color quality, aligned with CEE metric outcomes, affirming the effectiveness of the CLAHE method. Our findings represent a first step towards enhanced manuscript quality assessment. Future work is necessary to evaluate the improvement from the perspective of AI analysis of ancient manuscripts. The challenge for achieving this lies in the variability of the layout and font within manuscript images, which makes training for deep learning difficult. These advancements will contribute to a deeper understanding of the potential applications of contrast enhancement techniques

in digitized manuscript analysis and for digital preservation purposes.

6. Acknowledgements

We thank Dr. Alessia Cedola and Dr. Inna Bukreeva from the Institute of Nanotechnology (CNR) for their support and technical feedback on this research.

References

- [1] D. Hamidović, C. Clivaz, S. Bowen Savant, Ancient Manuscripts in Digital Culture: Visualisation, Data Mining, Communication (Volume 3), Brill, 2019.
- [2] A. Tonazzini, E. Salerno, Z. A. Abdel-Salam, M. A. Harith, L. Marras, A. Botto, B. Campanella, S. Legnaioli, S. Pagnotta, F. Poggialini, V. Palleschi, Analytical and mathematical methods for revealing hidden details in ancient manuscripts and paintings: A review, Journal of Advanced Research 17 (2019) 31–42. URL: https://www.sciencedirect.com/science/article/pii/S2090123219300037. doi:https://doi.org/10.1016/j.jare.2019.01.003, special Issue on Celebrating JAR-1st IF.
- [3] S. Brenner, R. Sablatnig, Subjective assessments of legibility in ancient manuscript images-the salami dataset, in: Pattern Recognition. ICPR International Workshops and Challenges: Virtual Event, January 10-15, 2021, Proceedings, Part VII, Springer, 2021, pp. 68–82.
- [4] L. P. Saxena, A survey of manuscripts digitization to restoration using image processing, imanager's Journal on Pattern Recognition 6 (2019) 27.
- [5] S. S. Agaian, K. Panetta, A. M. Grigoryan, Transform-based image enhancement algorithms with performance measure, IEEE Transactions on image processing 10 (2001) 367–382.
- [6] K. Panetta, C. Gao, S. Agaian, No reference color image contrast and quality measures, IEEE transactions on Consumer Electronics 59 (2013) 643–651.
- [7] S. S. Agaian, B. Silver, K. A. Panetta, Transform coefficient histogram-based image enhancement algorithms using contrast entropy, IEEE transactions on image processing 16 (2007) 741–758.
- [8] K. Panetta, Y. Zhou, S. Agaian, H. Jia, Nonlinear unsharp masking for mammogram enhancement, IEEE Transactions on Information Technology in Biomedicine 15 (2011) 918–928.
- [9] M. A. Qureshi, M. Deriche, A. Beghdadi, A. Amin, A critical survey of state-of-the-art image inpainting quality assessment metrics, Journal of Visual Communication and Image Representation 49 (2017) 177–191.
- [10] P. Mohammadi, A. Ebrahimi-Moghadam, S. Shirani, Subjective and objective quality assessment of image: A survey, arXiv preprint arXiv:1406.7799 (2014).
- [11] A. C. Bovik, Automatic prediction of perceptual image and video quality, Proceedings of the IEEE 101 (2013) 2008–2024.
- [12] M. Solh, G. AlRegib, Miqm: A multicamera image quality measure, IEEE transactions on image processing 21 (2012) 3902–3914.
- [13] Z. Sinno, A. C. Bovik, Large-scale study of perceptual video quality, IEEE Transactions on Image Processing 28 (2018) 612–627.
- [14] T. Tao, L. Ding, H. Huang, Unified non-uniform scale adaptive sampling model for quality assessment of natural scene and screen content images, Neurocomputing 399 (2020) 96–106.
- [15] L. Tang, C. Tian, L. Li, B. Hu, W. Yu, K. Xu, Perceptual quality assessment for multimodal medical image fusion, Signal Processing: Image Communication 85 (2020) 115852.
- [16] M. A. Qureshi, A. Beghdadi, M. Deriche, Towards the design of a consistent image contrast enhancement evaluation measure, Signal Processing: Image Communication 58 (2017) 212–227.
- [17] S. A. Amirshahi, A. Kadyrova, M. Pedersen, How do image quality metrics perform on contrast enhanced images?, in: 2019 8th European Workshop on Visual Information Processing (EUVIP), 2019, pp. 232–237. doi:10.1109/EUVIP47703.2019.8946143.

- [18] S.-D. Chen, A. R. Ramli, Minimum mean brightness error bi-histogram equalization in contrast enhancement, IEEE transactions on Consumer Electronics 49 (2003) 1310–1319.
- [19] Poem in honor of sultan sultans mehmed ii, by gian mario filelfo, Manuscript from Genève, Bibliothèque de Genève, Ms. lat. 99, p.1 - Amyris, Available at https://www.ecodices.ch/en/list/one/bge/lat0099 (2023/01/01), ????
- [20] R. Gonzalez, R. Woods, Digital Image Processing, Digital Image Processing, Prentice Hall, 2002. URL: https://books.google.it/books?id=YRRkQgAACAAJ.
- [21] J. V. Christian, Special issue on image quality assessment, in: Signal Process., vol. 64, no. 1, 1998, p. 129–130.
- [22] S. Paris, S. W. Hasinoff, J. Kautz, Local laplacian filters: Edge-aware image processing with a laplacian pyramid., ACM Trans. Graph. 30 (2011) 68.
- [23] D. M. Chandler, Seven challenges in image quality assessment: past, present, and future research, International Scholarly Research Notices 2013 (2013).
- [24] M. J. Melo, P. Nabais, M. Vieira, R. Araújo, V. Otero, J. Lopes, L. Martín, Between past and future: Advanced studies of ancient colours to safeguard cultural heritage and new sustainable applications, Dyes and Pigments 208 (2022) 110815.
- [25] A. Tonazzini, E. Salerno, Z. A. Abdel-Salam, M. A. Harith, L. Marras, A. Botto, B. Campanella, S. Legnaioli, S. Pagnotta, F. Poggialini, et al., Analytical and mathematical methods for revealing hidden details in ancient manuscripts and paintings: A review, Journal of advanced research 17 (2019) 31–42.
- [26] M. Hanif, A. Tonazzini, S. F. Hussain, U. Habib, E. Salerno, P. Savino, Z. Halim, Blind bleed-through removal in color ancient manuscripts, Multimedia Tools and Applications 82 (2023) 12321–12335.
- [27] I. Montani, E. Sapin, A. Pahud, P. Margot, Enhancement of writings on a damaged medieval manuscript using ultraviolet imaging, Journal of cultural heritage 13 (2012) 226–228.
- [28] J. Wang, Y. Hu, An improved enhancement algorithm based on cnn applicable for weak contrast images, IEEE Access 8 (2020) 8459–8476.
- [29] A. Mittal, A. K. Moorthy, A. C. Bovik, No-reference image quality assessment in the spatial domain, IEEE Transactions on image processing 21 (2012) 4695–4708.
- [30] G. Zhai, X. Min, Perceptual image quality assessment: a survey, Science China Information Sciences 63 (2020) 1–52.
- [31] J. Memon, M. Sami, R. A. Khan, M. Uddin, Handwritten optical character recognition (ocr): A comprehensive systematic literature review (slr), IEEE access 8 (2020) 142642–142668.
- [32] K. Matkovic, L. Neumann, A. Neumann, T. Psik, W. Purgathofer, et al., Global contrast factor-a new approach to image contrast., in: CAe, 2005, pp. 159–167.
- [33] M. Bressan, C. R. Dance, H. Poirier, D. Arregui, Local contrast enhancement, in: Color Imaging XII: Processing, Hardcopy, and Applications, volume 6493, SPIE, 2007, pp. 304–315.
- [34] C. Wang, Z. Ye, Brightness preserving histogram equalization with maximum entropy: a variational perspective, IEEE Transactions on Consumer Electronics 51 (2005) 1326–1334.
- [35] ITU-T.Recommendation, P.1301(07/2012), Subjective quality evaluation of audio and audiovisual multiparty telemeetings, Tech. rep., International Telecommunications Union, Geneva, Switzerland (July 2012).
- [36] ITU-T.Recommendation, P.910 (10/2023), Subjective video quality assessment methods for multimedia applications, Tech. rep., International Telecommunications Union, Geneva, Switzerland (Oct 2023).
- [37] e-codices virtual manuscript library of switzerland, university of fribourg, Available at *https://www.e-codices.ch* (2023/01/01), ????
- [38] Bible du xiiième siècle (first part: Genesis-psalms), Bern Burgerbibliothek Cod. 27, f. 2r Available at *https://www.e-codices.unifr.ch/it/list/one/bbb/0027* (2023/01/01), ????
- [39] Boethius, de consolatione philosophiae, Cologny Fondation Martin Bodmer Cod. Bodmer 41, f. 1r Available at *https://www.e-codices.ch/en/list/one/fmb/cb-0041* (2023/01/01), ????
- [40] Catullus, carmina, Cologny Fondation Martin Bodmer Cod. Bodmer 47, f. 1r Available at https://www.e-codices.unifr.ch/en/list/one/fmb/cb-0047 (2023/01/01), ????

- [41] Prudentius, carmina, Bern Burgerbibliothek Cod. 264, p. 61- Available at https://www.e-codices.unifr.ch/it/list/one/bbb/0264 (2023/01/01), ????
- [42] Divine comedy, Cologny, Fondation Martin Bodmer, Cod. Bodmer 56, f. 1r (Codice Ricasoli Firidolfi), Available at *https://www.e-codices.unifr.ch/en/list/one/fmb/cb-0056* (2023/01/01), ????
- [43] Gratianus, decretum (cum glossa ordinaria), Cologny Fondation Martin Bodmer Cod. Bodmer 75,
 f. 59r Available at *https://www.e-codices.unifr.ch/en/list/one/fmb/cb-0075* (2023/01/01), ????
- [44] Petrarca and dante, rhymes, Cologny Fondation Martin Bodmer Cod. Bodmer 131, f.8r Available at *https://www.e-codices.unifr.ch/en/list/one/fmb/cb-0131* (2023/01/01), ????
- [45] Ovid, metamorphoses, fasti, Cologny Fondation Martin Bodmer Cod. Bodmer 124, f. 2r Available at *https://www.e-codices.unifr.ch/en/list/one/fmb/cb-0124* (2023/01/01), ????
- [46] Prudentius, carmina, Bern Burgerbibliothek Cod. 264, p. 4- Available at https://www.ecodices.unifr.ch/it/list/one/bbb/0264 (2023/01/01), ????
- [47] Benevenutus imolensis, romuleon, Cologny Fondation Martin Bodmer Cod. Bodmer 143,f. 1r Available at *https://www.e-codices.unifr.ch/en/list/one/fmb/cb-0143* (2023/01/01), ????
- [48] Guiron le courtois, Cologny, Fondation Martin Bodmer, Cod. Bodmer 96-1, f. 1r Available at *https://www.e-codices.ch/en/list/one/bge/lat0096-1 (2023/01/01)*, ????
- [49] S. Long, X. He, C. Yao, Scene text detection and recognition: The deep learning era, International Journal of Computer Vision 129 (2021) 161–184.
- [50] C. K. Ch'ng, C. S. Chan, Total-text: A comprehensive dataset for scene text detection and recognition, in: 2017 14th IAPR international conference on document analysis and recognition (ICDAR), volume 1, IEEE, 2017, pp. 935–942.
- [51] Y. Liu, L. Jin, S. Zhang, C. Luo, S. Zhang, Curved scene text detection via transverse and longitudinal sequence connection, Pattern Recognition 90 (2019) 337–345.
- [52] Docsumo, Inc., Docsumo, https://docsumo.com, 2019. Accessed 2024.