Legibility vs. Extractability: Crafting Visual Defenses Against Automated OCR

Minh Nguyen*, Kien Tran1, The Han Huynh*

*Department of Computer and Information Sciences
Auckland University of Technology, Auckland, New Zealand

***Abstract*—The rise of generative AI, particularly large language models (LLMs), has redefined problem-solving across domains—but it has also introduced new challenges in academic integrity. Students are increasingly using AI-powered Optical Character Recognition (OCR) tools to extract restricted content from screenshots, bypassing traditional safeguards that prevent copying and pasting. In this study, we investigate a set of visual defense techniques designed to counter automated OCR systems: (1) adversarial fonts crafted to disrupt character recognition, (2) color-based distortions that alter visual contrast, (3) animated interference lines that obstruct character boundaries, and (4) a novel cloud blur effect that dynamically follows the cursor to obscure localized text regions. We evaluate these strategies across multiple LLM-integrated OCR platforms—ChatGPT, DeepSeek, Claude, and Gemini. Our findings show that modern OCR tools remain largely unaffected by custom fonts, line obstructions, and color distortions. In contrast, the cloud blur technique significantly reduces OCR accuracy while preserving legibility for human readers. These results highlight dynamic, context-aware visual obfuscation as a promising and potentially future-proof solution for deterring AI-assisted text extraction. Cloud blur, in particular, emerges as the most effective approach, offering strong resistance to OCR while maintaining accessibility for legitimate human users. A live demo is available at https://cv.aut.ac.nz/nFonts.**

***Index Terms*—Generative AI, Large Language Models, Optical Character Recognition, OCR Disruption, Custom Fonts, Visual Effects, AI in Education, Text Security.**

I. Introduction

The rapid proliferation of generative artificial intelligence (AI), particularly large language models (LLMs) [1], has transformed educational environments, providing unprecedented ease of access to instant answers and personalized learning support [2]. Despite these considerable benefits, generative AI has raised critical concerns regarding academic integrity, as students increasingly leverage AI-powered Optical Character Recognition (OCR) tools [3] to bypass conventional anti-cheating measures. While educational platforms often disable text copying to uphold academic standards, students routinely circumvent these safeguards by capturing screenshots and employing advanced OCR techniques [4]. The rise of LLMs is particularly relevant here, as their global adoption is accelerating rapidly—with projections estimating over 750 million applications integrating LLMs by 2025, and nearly 50% of all digital work expected to be automated using these tools [5]. These models, known for their flexibility and deep-learning capabilities, are also being used to enhance OCR performance, enabling more efficient extraction and interpretation of complex visual content.

The advancement of OCR technologies presents a substantial threat to the authenticity of educational assessments and learning materials. OCR-enabled AI tools are now capable of accurately extracting text from images, significantly complicating efforts to maintain assessment security and integrity. Thus, it is essential for educators and technologists to develop more robust strategies to combat unauthorized text extraction. In response to these challenges, this research introduces and evaluates three innovative methods specifically designed to impede AI-driven OCR technologies while preserving readability for human audiences. These methods include creating specialized OCR-resistant fonts, applying subtle visual effects, and incorporating dynamic visual disturbances. Firstly, custom-designed fonts are tailored to be challenging for OCR algorithms yet easy for human readers. Secondly, subtle visual modifications, such as carefully selected color patterns or textures, disrupt OCR processing while maintaining clear readability for users. Lastly, dynamic visual disturbances, including animated or randomly positioned interference lines, are introduced to further hinder AI-driven OCR extraction.

Research has demonstrated several effective techniques for disrupting optical character recognition (OCR) systems [6]. Some approaches involve the use of specially designed typefaces that alter letterforms to challenge text recognition. Additionally, studies have shown that visual perturbations, such as added noise or specific patterns, can significantly impact OCR performance while maintaining readability for human viewers. However, the rapid advancement in artificial intelligence, particularly in OCR technology integrated with modern AI platforms, has rendered many previous strategies less effective. Therefore, this research builds upon these foundational concepts by specifically examining the effectiveness of our proposed techniques against contemporary OCR-enabled AI tools [7], including ChatGPT, DeepSeek, Claude, and Gemini. By evaluating these advanced visual disruption methods, this study seeks to reinforce academic integrity and encourage genuine student engagement, striking a balance between leveraging technological innovation and maintaining ethical standards in educational environments.

II. Related Works

Generative AI and large language models (LLMs) have significantly transformed educational practices, presenting both novel learning opportunities and complex challenges concerning academic integrity. Barreto et al. (2023) investigated the dual implications of generative AI, noting its capacity for personalized learning alongside ethical and integrity concerns [2].

The evolution and accuracy improvements of Optical Character Recognition (OCR) technologies are thoroughly reviewed by Islam et al. (2017), highlighting their extensive applications in digitization and automated text processing across various sectors, including education [4]. However, the enhanced capability of OCR systems has increasingly compromised the security of educational assessments, driving the need for effective OCR-resistant strategies.

Adversarial fonts represent a significant approach in OCR disruption research. Mun's ZXX typeface explicitly challenges OCR by distorting conventional typographic forms while preserving human readability [8]. Further research has explored similar adversarial fonts and demonstrated their efficacy against traditional OCR algorithms through strategic distortions and modifications [9], [10].

Visual effect-based approaches have also been widely studied. Chen et al. (2020) demonstrated how carefully designed visual perturbations, including noise additions and pattern overlays, effectively degrade OCR accuracy without impairing human readability significantly [9]. Moreover, another research showed that subtle manipulations in color and texture can mislead advanced OCR algorithms, suggesting practical applications in protecting textual content from unauthorized extraction [11].

Dynamic visual disturbances constitute another promising strategy in OCR resistance. One highlighted the effectiveness of dynamic interference patterns, such as animated distortions, in obstructing contemporary OCR tools. These temporal disruptions create substantial challenges for OCR systems with minimal readability impact for humans [12].

The rapid advancements and integration of OCR within sophisticated generative AI platforms like ChatGPT, DeepSeek, Claude, and Gemini, underscore the need for ongoing innovation in OCR disruption methodologies. This research addresses existing gaps by systematically assessing the effectiveness of custom fonts, visual effects, and dynamic disturbances specifically against contemporary AI-integrated OCR platforms, contributing crucial insights toward maintaining academic integrity in AI-enhanced educational settings.

III. Proposed Approaches

To investigate practical defenses against AI-powered Optical Character Recognition (OCR) systems while maintaining acceptable human readability, we evaluated four visual techniques: (1) the use of customized fonts designed to challenge OCR text recognition, (2) the application of dynamic color distortions to disrupt feature consistency, (3) the overlay of animated interference lines to confuse OCR character boundaries, and (4) a dynamic "cloud blur" effect that introduces localized obfuscation based on cursor movement. Each method targets different stages of OCR pipelines such as character segmentation, feature extraction, and recognition. We then run a series of tests to evaluate which of these techniques most effectively disrupt OCR performance while remaining legible to human readers.

A. Font Creation for OCR Resistance

To systematically explore the limits of AI-based optical character recognition (OCR), we developed a series of custom fonts engineered to disrupt machine perception while preserving human readability. The fonts were created using FontForge and were designed to introduce increasingly complex visual transformations that mimic the ambiguity and variability found in handwritten scripts and non-Latin writing systems. These design interventions specifically target vulnerabilities in OCR systems that rely on consistent geometric features and standardized Unicode mappings.

1) Font Design Process: Using FontForge, we modified base Latin glyphs (a–z, A–Z) through a combination of manual editing and scripted transformations. The goal was to introduce structural anomalies without compromising the semantic recognizability of each character. The key techniques applied include:

- **Stroke Variation**: Non-uniform and asymmetric stroke thickness to obscure familiar character outlines and reduce geometric regularity.
- **Kerning Perturbation** (δ_s): Randomized inter-character spacing to disrupt alignment-based segmentation and character grouping.
- **Curvature Distortion** (κ): Altered arcs and corners that make shape matching and contour-based recognition less effective.
- **Character Substitution**: Selective replacement of Latin glyph segments with visual motifs derived from Chinese and Japanese scripts, introducing culturally unfamiliar structural elements that confuse models trained on Latin-centric datasets.

These transformations can be formalized as a parametric function:

$$g_i' = T(g_i, \lambda, \delta_s, \kappa) \tag{1}$$

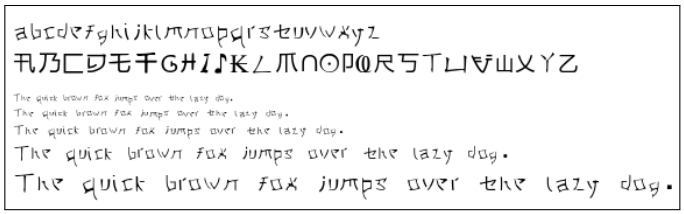
where g_i is the original glyph, T denotes the composite transformation function, λ controls stroke scaling, δ_s adjusts kerning, and κ governs curvature deformation. The output g_i' represents the distorted glyph used in the final font.

2) Font Variations: We created three progressively distorted fonts: `Font_1`, `Font_2`, and `Font_3`; each representing a different level of OCR resistance. All three fonts encode the full English alphabet in both uppercase and lowercase, along with standard punctuation, but differ in how radically they diverge from canonical Latin forms.

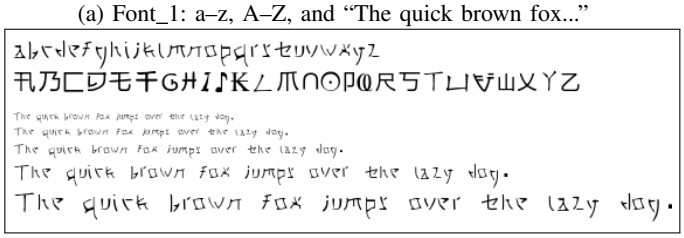
- **Font_1 (Moderate Modification)**: Maintains the overall structure of standard Latin characters but introduces

mild occlusions, irregular strokes, and subtle curvature changes. It is designed to be easily readable by humans but introduces enough noise to confuse basic OCR models.

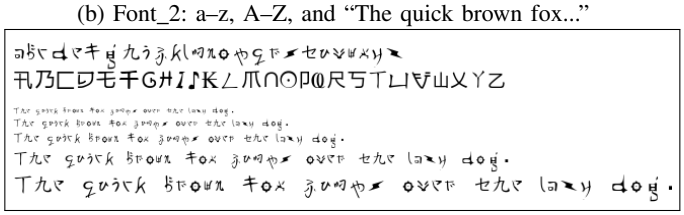
- **Font_2 (Intermediate Modification)**: Integrates pronounced features of East Asian scripts, such as dense stroke intersections, stroke imbalance, and partial glyph deconstruction. These designs remain interpretable to human readers with effort but significantly impair OCR performance.
- **Font_3 (Significant Modification)**: Features highly distorted and abstracted letterforms, with some glyphs bearing closer visual resemblance to logographic or symbolic forms. Human interpretation increasingly depends on sentence context, and OCR systems struggle to tokenize characters accurately.



(a) Font_1: a–z, A–Z, and "The quick brown fox..."



(b) Font_2: a–z, A–Z, and "The quick brown fox..."



(c) Font_3: a–z, A–Z, and "The quick brown fox..."

Fig. 1: Custom-designed fonts (Font_1, Font_2, Font_3), each showing full Latin alphabets and a sample sentence. Visual distortion increases from Font_1 to Font_3.

This progression enables a controlled experimental framework to evaluate OCR robustness. `Font_1` and `Font_2` are relatively readable, as the letters closely resemble standard English characters. In contrast, `Font_3` is significantly harder to read, as its letters appear similar to English but deviate noticeably, pushing the boundaries of legibility and typographic conventions. The use of culturally diverse typographic motifs serves not only to obfuscate machine perception but also to explore how cross-script features can be leveraged for information security and adversarial robustness.

B. Color Distortions to Impair OCR

Generally, OCR systems are optimized for high-contrast text—typically black characters on a white background. This assumption enables straightforward binarization and edge detection, which are foundational for accurate character segmentation. To challenge this assumption, we introduce dynamic color distortions by continuously varying the foreground (text), background, and gradient fill colors over time.

Unlike fixed color schemes, this method produces gradually and randomly evolving color combinations, including both high- and low-contrast moments. As shown in Figure 2, the text undergoes temporal transitions where, at certain moments, the foreground and background hues are poorly contrasted. These intervals may degrade OCR performance due to unreliable binarization. However, human visual perception—especially when aided by context and pattern recognition—can often tolerate these fluctuations, enabling readers to "fill in the gaps" even during low-contrast phases.

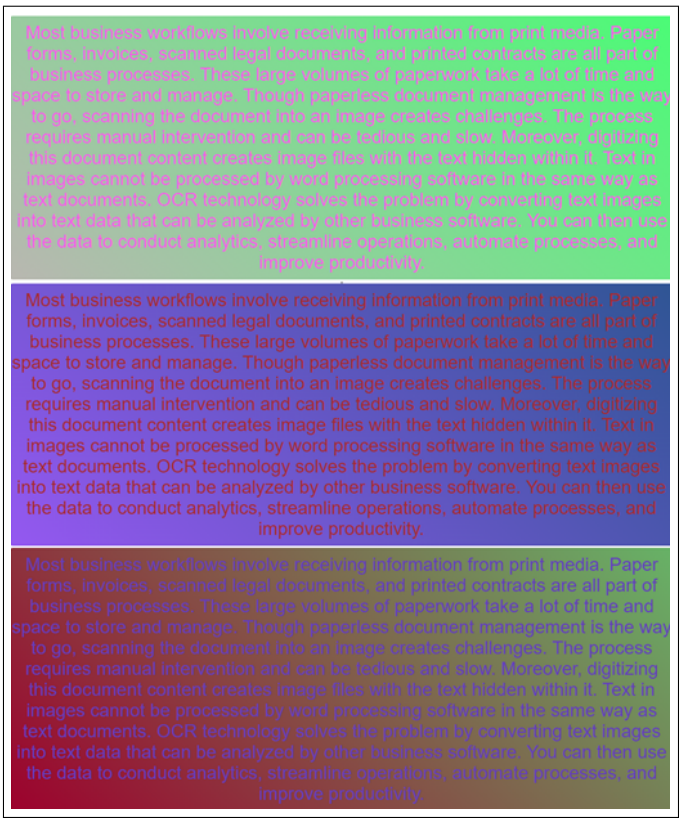


Fig. 2: Text displayed with gradually changing foreground and background colors across different time frames. While OCR may fail during low-contrast intervals, human perception compensates, preserving readability.

Formally, let $C_t = (R_t, G_t, B_t)$ represent the RGB vector of the text color at time t, and $B_t = (R_b, G_b, B_b)$ represent the background color vector. The instantaneous contrast between the text and background is given by:

$$D_t = \sum_{i \in \{R,G,B\}} |C_t^i - B_t^i| \tag{2}$$

To introduce continuous variation, we define color updates over time as:

$$C_{t+1} = C_t + \Delta C, \quad B_{t+1} = B_t + \Delta B \tag{3}$$

where ΔC and ΔB are small, randomized perturbation vectors ensuring smooth transitions in color space. Over time, this guarantees that:

$$\min_t D_t < \gamma \tag{4}$$

for some contrast threshold γ below which OCR systems are likely to struggle. The aim is to intermittently reach these low-contrast states while preserving an overall visual experience that remains interpretable to humans.

This dynamic coloring approach introduces temporal uncertainty for OCR systems. Traditional binarization and segmentation algorithms may fail to consistently extract contours when contrast is unstable or ambiguous. For humans, however, perceptual constancy and contextual inference allow comprehension to persist even when contrast momentarily degrades. This discrepancy raises an open question: *"How well can OCR models handle color dynamics that humans can tolerate?"* The following sections explore this question through empirical testing.

C. Dynamic Visual Disturbances via Color Moving Lines

OCR engines depend heavily on stable character alignment for accurate segmentation. To disrupt this stability, we introduce dynamic visual noise by overlaying interference lines that move continuously in both horizontal and vertical directions.

Let $L = l_1, l_2, \ldots, l_k$ denote a set of k moving lines, where each line l_i is characterized by its initial position (x_i, y_i) and a velocity vector (v_x, v_y). The position of each line at time t can thus be expressed as:

$$x_i(t) = x_i(0) + v_x t, \quad y_i(t) = y_i(0) + v_y t \tag{5}$$

where velocities v_x and v_y are randomly sampled from a uniform distribution to maintain variability:

$$v_x, v_y \sim U(-\alpha, \alpha) \tag{6}$$

By constraining α to small values, we preserve readability for humans while significantly impairing OCR segmentation algorithms. Humans rely on perceptual continuity and short-term memory, enabling comprehension even when text is momentarily obscured by moving lines. In contrast, OCR systems typically analyze static images; a single frame captured at any instant (such as the example shown in Figure 3a) contains fixed visual disturbances, severely reducing the accuracy of character recognition and segmentation.

D. Cloud Blur Effect with Mouse-Focused Visibility

To mimic the characteristics of human vision, where clarity is concentrated at the center of attention while peripheral details remain blurred; we introduce a dynamic cloud blur effect that only reveals sharp text around the mouse cursor.

Let $B(x, y)$ represent the blur intensity at pixel location (x, y), and let (x_m, y_m) denote the current cursor position. The blur field is defined as a Gaussian function centered on the cursor:

$$B(x, y) = \beta \left(1 - \exp\left(-\frac{(x - x_m)^2 + (y - y_m)^2}{2\sigma^2}\right)\right) \tag{7}$$

where β controls the maximum blur intensity and σ determines the spread of the clear region around the cursor.

This effect preserves readability within a focused region while rendering surrounding text visually degraded. Unlike conventional OCR, which requires uniformly sharp images for reliable text recognition, this technique disrupts recognition across most of the image. As demonstrated in Figure 3b, OCR systems are unable to segment or extract useful characters from the blurred background, while human users can easily interpret text by simply moving the cursor to the area of interest.

This visual strategy takes advantage of the fact that OCR systems cannot simulate spatial attention or cursor-based interaction, making it a practical defense mechanism against unauthorized text extraction.

E. Integration of Techniques

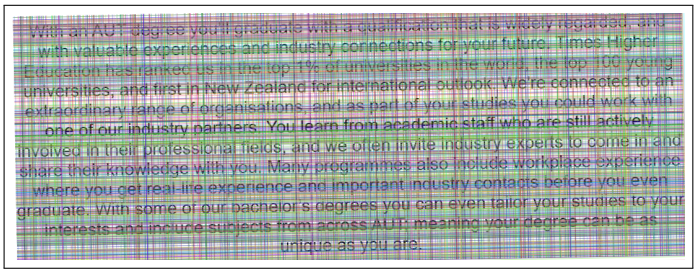
In this study, we have introduced four distinct techniques aimed at disrupting OCR-based text extraction: (1) OCR-resistant custom fonts, (2) bent color patterns with dynamic contrast, (3) random moving interference lines, and (4) cloud blur effects with cursor-focused clarity.

While each method individually impairs OCR performance, combining them can further strengthen the defense. However, excessive layering of these techniques also degrades human readability—defeating the purpose of accessible design.

Therefore, rather than compounding the effects, our focus is to evaluate each method independently to determine which offers the best balance between OCR resistance and human usability.

IV. Experimental Evaluation

To evaluate the effectiveness of our proposed OCR-resistant techniques, we conducted extensive experiments involving four widely-used large language models (LLMs): ChatGPT, DeepSeek, Claude, and Gemini. Each of these models includes robust AI-driven OCR capabilities, enabling them to analyze and interpret text directly from screenshots provided by users. The experiments involved presenting text images rendered in our specially designed fonts, further modified with dynamic visual effects. Performance was evaluated based on two primary criteria: the OCR recognition success rate of each model, and the readability of the text from a human perspective.



(a) Text with random moving colored lines. These animated lines obscure characters in any single static frame, confusing OCR systems. However, human vision is capable of filtering out dynamic noise through motion compensation, allowing readers to focus on and interpret the underlying text over time.

(b) Text with a cloud blur effect, where clarity is revealed only around the mouse cursor. This simulates human vision, which focuses sharply on a small area while the periphery remains blurred. The effect hinders OCR accuracy while preserving readability for users who can actively explore the content.

Fig. 3: Two OCR-disruptive visual techniques: (a) random color lines and (b) cloud blur effect. Both are designed to reduce machine readability while maintaining human legibility.

A. Evaluation Methodology

The evaluation consisted of two main components:

- **OCR Success Rate:** Measuring the ability of AI-based OCR tools to reconstruct the original text from processed images.
- **Human Readability Assessment:** Investigating the extent to which our techniques impact human legibility.

Each OCR tool analyzed text images under multiple conditions, including standard fonts, OCR-resistant fonts, and standard-font text with visual effects. The results help determine the most effective method for disrupting OCR while ensuring readability for human users.

B. OCR Success Rate Evaluation

To evaluate the effectiveness of OCR-resistant text styles, we tested several AI-based OCR tools on a series of modified text samples. Each tool was prompted to read and reproduce the content of various styled texts and report the percentage of correctly recognized characters. We then compared the outputs against the original input to calculate accuracy. In parallel, we assessed how readable these samples were to human readers.

1) Findings and Analysis: The results presented in Table I demonstrate that standard fonts, such as Arial, are consistently recognized with near-perfect accuracy ($\sim$100%) across all OCR tools evaluated. However, OCR-resistant fonts significantly reduce recognition performance, particularly for ChatGPT, which achieves only $\sim$20% accuracy on OCR-Resistant Font 1 and as low as $\sim$5% on OCR-Resistant Font 3. In contrast, the other models (DeepSeek, Claude, and Gemini) maintain high accuracy ($\sim$100%) on Fonts 1 and 2, though Gemini experiences a noticeable drop to $\sim$90% on Font 3. These findings suggest that font-based obfuscation alone may not be sufficient to consistently evade modern OCR systems, given their resilience to typographic variations.

Interestingly, not all visual effects degraded OCR performance as initially anticipated. Both color distortions (Effect 1) and moving lines (Effect 2) had minimal impact, with all tools maintaining approximately 90% recognition accuracy. This indicates that such effects are relatively ineffective at bypassing AI-based text extraction.

The most effective obfuscation was achieved using the cloud blur effect (Effect 3) applied to Arial text. All tools performed poorly under this condition, with success rates dropping to $\sim$5%. This suggests that even mild visual noise introduced through blurring can reliably hinder OCR without requiring exotic fonts or complex transformations.

2) Comparison with Human Readability: While OCR-resistant techniques can reduce recognition accuracy for AI models, their impact on human readability varies considerably. To evaluate this trade-off, a user study was conducted in which participants were asked to transcribe text rendered in various styles. Readability was rated on a five-point scale, from 1 (very difficult) to 5 (very easy).

As shown in Table I, standard fonts like Arial were rated as very easy to read. OCR-resistant fonts introduced moderate difficulty but remained generally understandable. More visually distorted styles, particularly OCR-Resistant Font 3, significantly impaired readability. Interestingly, some visual effects had little impact on human readers. Effects like cloud blur and moving lines further reduced legibility, though not as severely for humans as for AI models. These results suggest that certain visual effects can offer OCR resistance without substantially compromising human accessibility, especially when applied with care and moderation.

3) Implications: The findings have several important implications for the design of OCR-resistant content in contexts such as privacy protection, anti-scraping mechanisms, and secure document sharing. First, font-based obfuscation alone may not be sufficient to evade advanced OCR systems, as most models demonstrated high resilience against moderate typographic variation. This suggests that relying solely on font choice, even when using OCR-resistant designs, is unlikely to provide robust protection.

Second, the observed effectiveness of the cloud blur effect highlights the potential of simple visual noise as a practical defense mechanism. Since cloud blur consistently degraded OCR

TABLE I: OCR Success Rate Across Various AI Tools and Human Readability

Text Style	ChatGPT	DeepSeek	Claude	Gemini	Human Readability
Standard Font (Arial)	∼100%	∼100%	∼100%	∼100%	**Very Easy (5)**
OCR-Resistant Font 1	∼20%	∼100%	∼100%	∼100%	Moderate (3)
OCR-Resistant Font 2	∼15%	∼100%	∼100%	∼100%	Moderate (3)
OCR-Resistant Font 3	∼5%	∼5%	∼5%	∼90%	Difficult (2)
Arial with Color Distortions (Effect 1)	∼90%	∼90%	∼90%	∼90%	Moderate (3)
Arial with Moving Lines (Effect 2)	**∼0%**	∼90%	∼90%	∼90%	Moderate (3)
Arial with Cloud Blur (Effect 3)	∼5%	**∼5%**	**∼5%**	**∼5%**	Easy (4)

success rates across all models while maintaining reasonable human readability, it presents a viable option for protecting sensitive textual information from automated extraction.

Finally, the minimal impact of effects like color distortion and moving lines on both OCR performance and human readability suggests they are less effective on their own but could be valuable when used in combination with other strategies. For instance, layering modest visual effects with lightly altered fonts may increase resistance without compromising usability.

V. Discussion and Conclusion

This study explored the growing tension between legibility and extractability in the era of AI-powered OCR systems. With the increasing integration of large language models (LLMs) into OCR pipelines, traditional strategies for preventing unauthorized text extraction are becoming less effective. Our investigation focused on evaluating four key techniques: custom OCR-resistant fonts, color distortions, dynamic interference lines, and a novel cloud blur effect, designed to impede OCR performance while preserving human readability.

The results indicate that while adversarial fonts can offer moderate protection, most modern OCR-enabled AI tools are robust against typical typographic variations. Notably, only ChatGPT showed significant performance degradation with the OCR-Resistant fonts we made, whereas other tools such as DeepSeek, Claude, and Gemini maintained near-perfect accuracy for the same fonts.

Visual effects, on the other hand, demonstrated a more nuanced impact. Color distortions and moving lines offered limited resistance to OCR systems but preserved high levels of human readability. These techniques may serve best as complementary layers in a multi-faceted defense strategy. The most impactful and promising technique was the cloud blur effect, which achieved substantial OCR disruption across all tools tested, reducing recognition rates to approximately 5% while remaining easily readable to human users through cursor-based interaction.

From a practical perspective, these findings point toward cloud blur as a viable, future-proof defense mechanism for safeguarding visual content from automated OCR. Its ability to balance machine resistance with human accessibility makes it suitable for use in educational settings, secure content sharing platforms, and anti-scraping scenarios.

We strongly believe that even as OCR systems continue to advance in precision and adaptability, the cloud blur technique will remain effective. Because it introduces dynamic, user-dependent clarity that cannot be captured in a single static frame—on which OCR systems rely—it inherently resists conventional and AI-enhanced recognition techniques. As such, it offers a robust and sustainable defense, adaptable to future developments in machine vision.

Looking ahead, further research could explore adaptive or context-aware obfuscation strategies that dynamically respond to the capabilities of specific OCR systems. Additionally, integrating user feedback and eye-tracking data could help fine-tune visual effects to optimize human readability.

Ultimately, as AI-powered OCR tools continue to evolve, so too must our defenses. This study contributes a step in that direction by identifying promising techniques that do not rely solely on obscurity or excessive distortion but instead leverage perceptual insights and interactive elements to protect text in a visually intelligent way.

References

[1] A. A. Linkon, M. Shaima, M. S. U. Sarker, B. Badruddowza, N. Nabi, M. N. U. Rana, S. K. Ghosh, M. A. Rahman, H. Esa, and F. R. Chowdhury, "Advancements and applications of generative artificial intelligence and large language models on business management: A comprehensive review," *Journal of Computer Science and Technology Studies*, vol. 6, no. 1, pp. 225–232, 2024.

[2] F. Barreto, L. Moharkar, M. Shirodkar, V. Sarode, S. Gonsalves, and A. Johns, "Generative artificial intelligence: Opportunities and challenges of large language models," in *International conference on intelligent computing and networking*. Springer, 2023, pp. 545–553.

[3] X. Zhang, W. Bai, and H. Cui, "Review of optical character recognition for power system image based on artificial intelligence algorithm," *Energy Engineering*, vol. 120, no. 3, pp. 665–679, 2023.

[4] N. Islam, Z. Islam, and N. Noor, "A survey on optical character recognition system," *arXiv preprint arXiv:1710.05703*, 2017.

[5] K. Hooda and A. Mansukhani, "50+ essential llm usage stats you need to know in 2025," December 2024, accessed: 2025-03-27. [Online]. Available: https://keywordseverywhere.com/blog/llm-usage-stats/

[6] J. Deng, L. Dong, J. Chen, D. Yan, R. Wang, D. Ye, L. Zhao, and J. Tian, "Universal defensive underpainting patch: Making your text invisible to optical character recognition," in *Proceedings of the 31st ACM International Conference on Multimedia*, 2023, pp. 7559–7568.

[7] T. Gao, J. Jin, Z. T. Ke, and G. Moryoussef, "A comparison of deepseek and other llms," *arXiv preprint arXiv:2502.03688*, 2025.

[8] S. Mun, "Zxx: A typeface to thwart ocr recognition," 2012, zXX is a disruptive typeface designed to be unreadable by OCR software, featuring six different styles: Sans, Bold, Camo, False, Noise, and Xed. [Online]. Available: https://zxxfont.com

[9] L. Chen and W. Xu, "Attacking optical character recognition (ocr) systems with adversarial watermarks," *arXiv preprint arXiv:2002.03095*, 2020.

[10] Z. Shen, H. Luo, S. Li, and T. Li, "Adversarial training with ocr modality perturbation for scene-text visual question answering," in *2024 IEEE International Conference on Multimedia and Expo (ICME)*. IEEE, 2024, pp. 1–6.

[11] Y. He, K. Chen, G. Chen, Z. Ma, K. Zhang, J. Zhang, H. Bian, H. Fang, W. Zhang, and N. Yu, "Protego: Protect text content against ocr extraction attack," in *Proceedings of the 31st ACM International Conference on Multimedia*, 2023, pp. 7424–7434.

[12] M. Tayara, "The robustness of animated text captchas," Ph.D. dissertation, Newcastle University, 2017.