

Handling Strategies for Artificial Intelligence Misuse via Deepfakes in Digital Information Spreading from an Asymmetric Warfare Perspective

Pasha Atta Usman¹

¹Indonesia Defense University, Salemba, Jakarta, Indonesia

F.G. Cempaka Timur²

²Indonesia Defense University, Salemba, Jakarta, Indonesia

Pujo Widodo³

³Indonesia Defense University, Salemba, Jakarta, Indonesia

*Corresponding Author: pshatta08@gmail.com

Abstract

The rapid advancement of Artificial Intelligence (AI) has led to new security vulnerabilities, particularly through the malicious deployment of deepfakes in digital information dissemination. During the civil demonstrations of August 2025 in Indonesia, deepfake technology was aggressively utilized as a tool of asymmetric warfare to circulate institutional hoaxes, defame security personnel, and trigger instantaneous public unrest. This study aims to evaluate the strategic handling and mitigation countermeasures implemented by national security agencies specifically BSSN, Komdigi, Kemhan, and BIK Intelhan in neutralizing these cognitive threats. Utilizing a qualitative descriptive method with a comprehensive literature and policy review, this research analyzes the operational capabilities and institutional gaps within Indonesia's cyber defense architecture. The findings indicate that while primary detection frameworks exist, a highly integrated, real-time counter-propaganda mechanism and synchronized cross-sectoral coordination are still heavily required to counter modern asymmetric digital operations. Consequently, this study provides strategic policy recommendations to fortify national stability and digital resilience against sophisticated synthetic media operations.

Keywords: Artificial Intelligence, Deepfakes, Handling Strategies, Asymmetric Warfare, August 2025 Demonstrations.

1591

IJEVSS, Volume 05 Issue 04, 2026



Copyright at authors some right reserved this work is licensed under a [Creative Commons Attribution-Share Alike 4.0 International License](https://creativecommons.org/licenses/by-sa/4.0/)

INTRODUCTION

The rapid advancement of Artificial Intelligence (AI) technology has introduced unprecedented challenges to global security architectures (Giles, 2020). While AI offers significant technological breakthroughs, its dual-use nature has been increasingly exploited for malicious activities, particularly through the creation and dissemination of deepfakes (Chesney & Citron, 2019). In the contemporary strategic landscape, the manipulation of digital information using hyper-realistic synthetic media is no longer merely a cybercrime issue; rather, it has transformed into a sophisticated tool of asymmetric warfare (Blank, 2021). This method strategically targets the cognitive domain of a society, exploiting political polarization and social vulnerabilities to destabilize state institutions without relying on conventional military force (Waltzman, 2018). Consequently, developing robust handling strategies against such modern threats has become imperative for national defense agencies worldwide.

The gravity of this threat was vividly demonstrated during the civil unrest in Indonesia in August 2025. During these demonstrations, malicious actors extensively utilized deepfake technology as an offensive instrument in the digital information sphere (National Cyber and Crypto Agency [BSSN], 2025). Hyper-realistic manipulated videos of government officials were circulated across social media platforms to alter public perceptions, alongside synthetically generated audio and video designed to defame security forces. This systematic dissemination of deepfake-driven hoaxes succeeded in instigating instant public outrage, accelerating mass mobilization, and complicating the state's crowd control efforts. Within the framework of asymmetric warfare, this phenomenon represents a tactical cognitive strike, where psychological manipulation is synchronized with digital subversion to exploit domestic political instability (Lasconjarias & Larsen, 2020).

Addressing such sophisticated digital threats requires a well-coordinated national security response. In the Indonesian context, managing the fallout of the August 2025 incidents involves multiple critical stakeholders, including the National Cyber and Crypto Agency (BSSN), the Ministry of Communication and Digital Affairs (Komdigi), the Ministry of Defense (Kemhan), and the Defense Intelligence Agency (BIK Intelhan). While these institutions possess specialized mandates in cyber defense and strategic intelligence, the rapid evolution of deepfakes has exposed operational gaps in early detection, real-time counter-propaganda, and cross-sectoral coordination (Ministry of Defense [Kemhan], 2025). Therefore, analyzing handling strategies from an asymmetric warfare perspective is essential to build an integrated defense mechanism capable



of neutralizing hostile cognitive operations before they escalate into national security crises (Saronto, 2022).

While previous studies on deepfakes have predominantly focused on technical detection algorithms or general legal frameworks (Verdoliva, 2020), there remains a critical research gap in analyzing these synthetic media threats through the lens of national defense strategy, specifically within an asymmetric warfare context. This research addresses this gap by providing a comprehensive analysis of the multi-agency response during the August 2025 crisis. The primary objective of this study is to formulate strategic handling recommendations that enhance cross-sectoral coordination among BSSN, Komdigi, Kemhan, and BIK Intelhan. By evaluating institutional vulnerabilities and operational bottlenecks, this paper proposes an integrated cyber-cognitive defense model and policy improvements necessary for early detection and rapid counter-propaganda. Ultimately, this study contributes to the conceptual development of modern asymmetric warfare countermeasures, ensuring the state's resilience against sophisticated digital subversion.

RESEARCH METHODS

This study employs a qualitative descriptive research method to analyze the strategic countermeasures against deepfake misuse within the framework of national security. According to established academic methodologies, a qualitative descriptive approach is uniquely suited for investigating complex, contemporary phenomena within their real-life contexts, particularly where the boundaries between the phenomenon and the context are highly integrated (Creswell & Creswell, 2018; Yin, 2018). The primary data and information for this research are gathered through a comprehensive literature study and document analysis, often termed as desk research (Snyder, 2019).

Secondary data sources include official government regulations, national security policy documents, strategic defense white papers, academic journals, and verifiable media archives documenting the precise timeline of the August 2025 demonstrations in Indonesia. To ensure analytical rigor, the collected data undergoes a thematic policy evaluation, cross-referencing the operational and tactical responses of BSSN, Komdigi, Kemhan, and BIK Intelhan against established theories of asymmetric and cognitive warfare (Saronto, 2022). This methodological design aims to evaluate institutional capability gaps and produce structured, strategic policy recommendations for modern hybrid threat mitigation.



RESULT AND DISCUSSION

The Mechanism of Deepfake Misuse in Asymmetric Warfare

In the era of modern strategic conflict, asymmetric warfare has shifted from physical battlefields to digital and cognitive domains (Blank, 2021). The utilization of Artificial Intelligence (AI) via deepfakes represents a significant evolution in hybrid threats, where hyper-realistic synthetic media is deployed to destabilize an adversary's domestic security without engaging in conventional military actions (Waltzman, 2018). During the civil unrest in Indonesia in August 2025, the mechanism of deepfake proliferation followed a highly calculated, multi-staged offensive pattern designed to exploit existing societal vulnerabilities.

The tactical execution of this digital operation relied on three primary mechanisms: high-fidelity visual manipulation, rapid cross-platform amplification, and psychological synchronization. First, malicious actors used deep learning algorithms to generate synthetic video and audio depicting top government officials making controversial statements and security personnel executing unauthorized, violent suppressions. Second, these deepfakes were deliberately introduced into high-velocity social media platforms, such as TikTok and X (formerly Twitter), where algorithmic recommendation engines accelerated their reach before official verification could take place. Third, these operations were synchronized with existing political grievances, successfully triggering instantaneous public outrage and mass mobilization. Within asymmetric warfare theory, this mechanism is classified as a cognitive strike—a method that weaponizes information to paralyze the state's crisis response and undermine public trust in national defense and governance institutions (Lasconjarias & Larsen, 2020).

Institutional Responses and Operational Gaps during the August 2025 Crisis

The digital escalation during the August 2025 demonstrations posed a critical test for Indonesia's cyber defense and intelligence architecture. Mitigating the spread of deepfake-driven disinformation required an immediate, synchronized response from key national security stakeholders, namely the National Cyber and Crypto Agency (BSSN), the Ministry of Communication and Digital Affairs (Komdigi), the Ministry of Defense (Kemhan), and the Defense Intelligence Agency (BIK Intelhan). However, the operational reality during the crisis exposed several structural vulnerabilities and coordination bottlenecks within the state's response mechanism (Ministry of Defense [Kemhan], 2025).

The primary operational gap identified during the crisis was the absence of a real-time, unified detection and counter-propaganda framework. While BSSN focused on technical cyber



monitoring and security infrastructure, and Komdigi handled the downstream takedown of malicious content on public social media, a significant latency occurred between the initial detection of a deepfake and its actual removal. This response lag allowed the synthetic media to achieve maximum viral penetration and cognitive impact before the public could be informed of its fraudulent nature. Furthermore, from a defense intelligence perspective, BIK Intelhan faced challenges in rapidly synchronizing tactical cyber intelligence with operational public communications. This institutional compartmentalization—often exacerbated by bureaucratic silos and overlapping jurisdictions—hampered early warning capabilities and prevented the execution of a pre-emptive strategic narrative. Consequently, the crisis underscored that conventional, fragmented cyber security protocols are highly inadequate for countering synchronized asymmetric cognitive operations (National Cyber and Crypto Agency [BSSN], 2025; Saronto, 2022).

Strategic Handling Recommendations: Formulating an Integrated Cyber-Cognitive Defense

To neutralize the evolving threat of deepfakes in asymmetric warfare, Indonesia must transition from a reactive, fragmented cyber security posture to a proactive, integrated cyber-cognitive defense architecture (Blank, 2021). The operational bottlenecks observed during the August 2025 crisis demonstrate that technological detection alone is insufficient; it must be coupled with rapid institutional synchronization and strategic narrative management (Waltzman, 2018). Therefore, a structured handling strategy requires the formal establishment of a Joint Cyber-Cognitive Task Force, legally and operationally uniting BSSN, Komdigi, Kemhan, and BIK Intelhan into a single command framework.

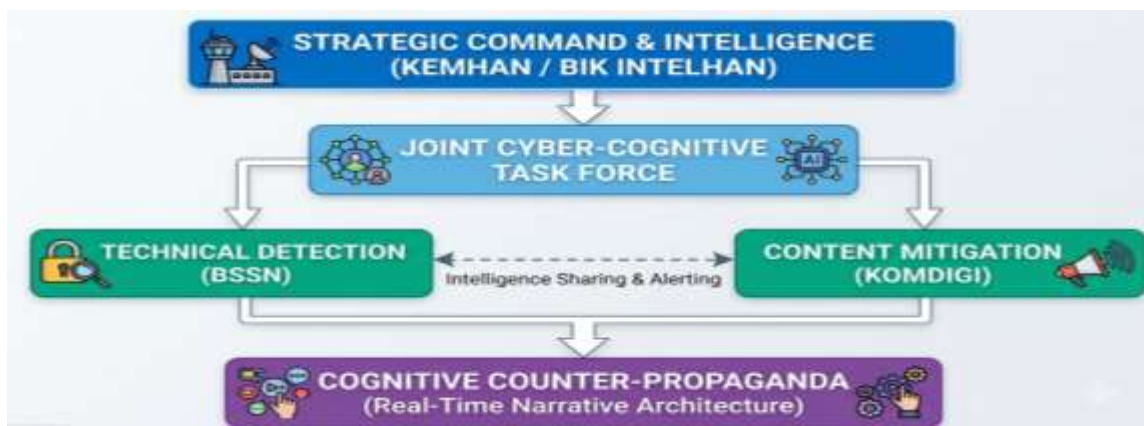


Figure 1. The Integrated Cyber-Cognitive Defense Framework.

This integrated model operates through three synchronized pillars: automated early warning, streamlined inter-agency protocols, and active cognitive resilience. First, BSSN and BIK Intelhan must deploy advanced AI-driven deepfake detection algorithms at the structural gateway level to intercept malicious synthetic media before it achieves viral velocity (Verdoliva, 2020). Second, once a digital threat is identified, an automated protocol must immediately trigger Komdigi for rapid downstream mitigation and content labeling, effectively cutting through bureaucratic delays. Third, Kemhan must spearhead the public-facing response by deploying real-time counter-propaganda and authoritative strategic narratives to pre-emptively neutralize public panic (Lasconjarias & Larsen, 2020). By institutionalizing this collaborative defense mechanism, the state can effectively safeguard its information ecosystem, protect public order, and defeat sophisticated cognitive operations designed to exploit national vulnerabilities (Saronto, 2022).

CONCLUSION

The weaponization of Artificial Intelligence via deepfakes during the August 2025 demonstrations represents a significant shift in asymmetric warfare tactics within Indonesia's strategic environment. This study concludes that the state's traditional, fragmented cyber defense framework is highly inadequate against synchronized cognitive subversion designed to exploit institutional latency and societal vulnerabilities. While critical agencies—specifically BSSN, Komdigi, Kemhan, and BIK Intelhan—possess distinct technical competencies, their operational responses were heavily constrained by bureaucratic silos and the lack of a real-time, unified protocol. To defeat modern hybrid threats, this paper strongly recommends transitioning toward an integrated cyber-cognitive defense architecture through the institutionalization of a Joint Task Force. By establishing automated multi-agency early warning systems and dynamic counter-propaganda mechanisms, Indonesia can effectively protect its digital information sphere, safeguard public trust, and preserve domestic national security in the face of rapid technological disruption.

REFERENCES

- Blank, S. (2021). *Cyber and cognitive threats in modern asymmetric conflict*. Strategic Studies Institute.
- Chesney, B., & Citron, D. (2019). Deepfakes and the new disinformation economy. *California Law Review*, 107(6), 1753–1820.



- Creswell, J. W., & Creswell, J. D. (2018). Research design: Qualitative, quantitative, and mixed methods approaches (5th ed.). SAGE Publications.
- Giles, K. (2020). The next phase of Russian information warfare. NATO Defense College.
- Lasconjarias, G., & Larsen, J. A. (Eds.). (2020). NATO's response to hybrid threats. NATO Defense College.
- Ministry of Defense [Kemhan]. (2025). Strategic evaluation on cyber security and information dissemination during national instability. Indonesian Ministry of Defense.
- National Cyber and Crypto Agency [BSSN]. (2025). Annual report on synthetic media threats and cognitive operations in Indonesia. BSSN Press.
- Saronto, W. (2022). Pembangunan jaringan. ANDI Yogyakarta.
- Snyder, H. (2019). Literature review as a research methodology: An overview and guidelines. *Journal of Business Research*, 104, 333–339. <https://doi.org/10.1016/j.jbusres.2019.07.039>
- Verdoliva, L. (2020). Media forensics for the detection of deepfakes: A survey. *IEEE Access*, 8, 97932–97963. <https://doi.org/10.1109/ACCESS.2020.2997159>
- Waltzman, R. (2018). The weaponization of information: The need for cognitive security. RAND Corporation.
- Yin, R. K. (2018). Case study research and applications: Design and methods (6th ed.). SAGE Publications.

