THE RISING THREAT OF DEEPFAKES: IS MALAYSIA DOING ENOUGH TO STOP THEM?

Agnes Lim Hai Thong*

Abstract

The rapid evolution of Artificial Intelligence ("AI") and Deepfake technology has outpaced Malaysia's existing legal and regulatory frameworks, exposing critical vulnerabilities in addressing their misuse. Their sophisticated capabilities pose significant threats, extending beyond personal victimization to national security concerns. Malaysia's reliance on existing fragmented legal provisions, instead of creating AI-specific legislation, may adequately address most categories of deepfake-related offenses despite their broad application. Nevertheless, these provisions are predominantly reactive and criminal in nature, targeting perpetrators while leaving various loopholes in tackling deepfake crimes. These loopholes include the requirements for holding perpetrators liable, such as proving the malicious intention to misuse fake content, the existence of actual harm to victims, the difficulty of tracing the true identity of perpetrators, and the requirement that the content must be generated by AI with demonstrable negative impacts on victims. In addition, the current legal framework indicates that victims face substantial barriers in pursuing civil remedies, leaving them uncompensated for their suffering. The author of this paper urges for comprehensive reform of Malaysia's legal framework on civil remedies and proactive regulatory measures, along with the development of AI detection tools to keep pace with increasingly sophisticated deepfakes, and the promotion of public awareness of deepfake threats through enhanced media literacy education. With these suggestions, society would be better protected from deepfake threats.

* LLB (Honours) candidate under the HULLB programme, HELP University.

1. Introduction

The aim of this paper is to examine and evaluate Malaysia's existing regulatory response to the proliferation of deepfake technology and its associated risks. The global accessibility of internet resources enables the rapid dissemination of technological developments, including sophisticated deepfake creation tools.[1] This implies that the creation of deepfakes is becoming increasingly easy and inexpensive while maintaining near-undetectable quality. Most jurisdictions have not specifically criminalized deepfake technology, recognizing its legitimate applications and transformative potential across various sectors.[2] However, this regulatory restraint has inadvertently facilitated the technology's exploitation for fraudulent purposes, creating substantial societal risks. In Malaysia, the proliferation of deepfake-related incidents, coupled with inadequate regulatory frameworks, demonstrates the inability of existing legal mechanisms to protect citizens from technological abuse. Therefore, this paper not only explains what a deepfake is and how it works, but also highlights its dangerously far-reaching potential and discusses Malaysia's current legal framework. Finally, the author suggests comprehensive legal reforms and additional measures to address the ineffectiveness of the current framework in tackling deepfake threats.

2. Define Deepfake Technology

What exactly is a deepfake? Deepfake technology derives its nomenclature from the combination of "deep learning" and "fake production," reflecting the AI methodology employed in content manipulation.[3] Specifically, deepfakes constitute AI-generated content in which an individual's facial features and vocal characteristics are digitally substituted with those of another person,[4]

---

[1] Tan Zec Kie, Chong Shao Zheng, Kuek Chee Ying and Tay Eng Siang, 'Individual Legal Protection in the Deepfake Technology Era' (2023) Proceedings of the 3rd International Conference on Law and Digitalization 2023 <https://www.atlantis-press.com/article/125995062.pdf> accessed 20 July 2025.

[2] Nicholas Caporusso, 'Deepfakes for the Good: a Beneficial Application of Contentious Artificial Intelligence Technology' (2021) Advances in Artificial Intelligence, Software and Systems Engineering <https://www.researchgate.net/publication/342691593_Deepfakes_for_the_Good_A_Beneficial_Application_of_Contentious_Artificial_Intelligence_Technology> accessed 25 July 2025.

[3] Tan Zec Kie, Chong Shao Zheng, Kuek Chee Ying and Tay Eng Siang, 'Individual Legal Protection in the Deepfake Technology Era' (2023) Proceedings of the 3rd International Conference on Law and Digitalization 2023 <https://www.atlantis-press.com/article/125995062.pdf> accessed 20 July 2025.

[4] Ángel Vizoso, Martín Vaz-Álvarez, and Xosé López-García, 'Fighting Deepfakes: Media and Internet Giants' Converging and Diverging Strategies Against Hi-Tech Misinformation' (2021) 9(1) Media and Communication 291–300. <https://doi.org/10.17645/mac.v9i1.3494> accessed 26 June 2025.

enabling the creation of fabricated content featuring the target individual without their consent or knowledge. One might assume that deepfake content is easy to detect, as it may contain unusual noises, mismatched facial expressions and voices, or even glitches.[5] However, this technology has now achieved unprecedented levels of sophistication,[6] rendering detection increasingly challenging, even for experts, without the aid of supporting tools.

3. The Development of Deepfake Technology

The Creation of A Deepfake

Deepfake creation employs sophisticated multi-stage processes that exceed public understanding of the technology's complexity.[7] It is essential to understand these stages to fully appreciate the unimaginably negative potential that deepfake technology possesses.

The generation process encompasses five distinct phases:

1. Data collection, wherein AI systems aggregate source material from publicly accessible platforms, including social media networks and search engines;
2. Initial content generation through algorithmic processing of collected data;
3. Comparative analysis between generated and authentic content to identify any discrepancies;
4. Iterative refinement based on analytical feedback to enhance realism; and
5. Final validation to ensure near-indistinguishable quality from genuine content.

The technological barriers to deepfake creation continue to diminish as computational tools become increasingly accessible and cost-effective.

---

[5] Anis Zalani, 'AI scams are getting real: Here are the cases happening in Malaysia that you should know about' *Malaymail* (Malaysia, 4 August 2025) <https://www.malaymail.com/news/malaysia/2025/08/04/ai-scams-are-getting-real-here-are-the-cases-happening-in-malaysia-that-you-should-know-about/183459> accessed 19 August 2025.

[6] Nicholas Caporusso, 'Deepfakes for the Good: a Beneficial Application of Contentious Artificial Intelligence Technology' (2021) Advances in Artificial Intelligence, Software and Systems Engineering <https://www.researchgate.net/publication/342691593_Deepfakes_for_the_Good_A_Beneficial_Application_of_Contentious_Artificial_Intelligence_Technology> accessed 25 July 2025.

[7] Gabe Regan, 'How Deepfakes Are Made' (*Startup Defender*, 16 June 2025) <https://www.realitydefender.com/insights/how-deepfakes-are-made> accessed 20 August 2025.

4. The Danger of Deepfake

With the constant advancement of AI tools, fake content may appear so hyper-realistic that it is nearly indistinguishable from genuine material. Such content is often created featuring publicly known individuals to spread misinformation to the general public, whether with good or bad intentions, or even unintentionally, in order to influence views on particular subjects for indirect benefit. However, as deepfake use has become widely publicized, perpetrators are now motivated by greed beyond simply spreading misinformation. This makes the existence and ability to create deepfakes appear even more sinister than we might originally think.

In today's digital world, disinformation is what perpetrators aim for.[8] They utilize deepfake technology to intentionally deceive the public, often in exchange for money or personal data.[9] Targets may include an individual, a group, a country, or even the global community. The impact of a single act is unimaginable; it could substantially disrupt an individual's life or even destabilize a nation. At the individual level, disruption caused by deepfakes may include financial loss, reputational damage, or even harm to one's mental health. At the societal level, destabilization could occur through AI-generated content, particularly when created in the name of the government or prominent politicians highly respected or relied upon by the public, that fabricates facts designed to incite negative emotions and erode trust.

In Malaysia, over the past two years, numerous scam cases have involved victims receiving phone calls that sounded exactly like their close relatives, requesting urgent financial help. These audios were later discovered to be AI-generated using voice samples collected from WhatsApp voice notes, Instagram stories, or public videos. Such cases succeed by creating emotional pressure, triggering immediate reactions from the victims before their brain can process what is happening.[10] In such emotionally charged situations, people often believe what they see or hear,

[8] Anis Zalani, 'That voice isn't real: Why media literacy matters in the age of deepfake and AI scams' *Malaymail* (Malaysia, 5 August 2025) <https://www.malaymail.com/news/malaysia/2025/08/05/that-voice-isnt-real-why-media-literacy-matters-in-the-age-of-deepfake-and-ai-scams/183565> accessed 19 August 2025.

[9] Anis Zalani, 'That voice isn't real: Why media literacy matters in the age of deepfake and AI scams' *Malaymail* (Malaysia, 5 August 2025) <https://www.malaymail.com/news/malaysia/2025/08/05/that-voice-isnt-real-why-media-literacy-matters-in-the-age-of-deepfake-and-ai-scams/183565> accessed 19 August 2025.

[10] Anis Zalani, 'That voice isn't real: Why media literacy matters in the age of deepfake and AI scams' *Malaymail* (Malaysia, 5 August 2025) <https://www.malaymail.com/news/malaysia/2025/08/05/that-voice-isnt-real-why-media-literacy-matters-in-the-age-of-deepfake-and-ai-scams/183565> accessed 19 August 2025.

especially when the deepfake is highly convincing, instead of questioning whether the content could be forged.

In 2024, the Director of Bukit Aman's Commercial Crime Investigation Department (CCID), Datuk Seri Ramli Mohamed Yoosuf, reported 454 fraud cases involving deepfake technology, with total losses amounting to RM2.72 million.[11] This demonstrates that many people continue to fall for fake content and that AI has advanced to a near-indistinguishable level.

However, deepfakes present risks beyond monetary loss, as they can divide society or influence public opinion. Besides deepfake phone calls, scammers have leveraged AI to produce videos featuring well-known figures, exploiting their reputations to trick victims into bogus investment schemes. Very recently, the Head of the MCA's Public Services and Complaints Department, Datuk Seri Michael Chong, revealed that Malaysians lost RM2.11 billion to scams last year, with 13,956 cases reported.[12] Notably, 85% of victims were convinced to invest after watching fake promotional videos seemingly endorsed by public figures.[13]

One recent example is a TikTok video featuring the Malaysian Border Control and Protection Agency's Director-General, Shuhaily Zain, promoting a service provider for individuals facing passport and official document issues.[14] The video, which was AI-generated, aimed to smear Shuhaily's image and mislead the public. Although the clip was uploaded by "Hermano Kedutaan," no further updates were available regarding whether the true perpetrator had been identified and penalized.

[11] Anis Zalani, 'AI scams are getting real: Here are the cases happening in Malaysia that you should know about' *Malaymail* (Malaysia, 4 August 2025) <https://www.malaymail.com/news/malaysia/2025/08/04/ai-scams-are-getting-real-here-are-the-cases-happening-in-malaysia-that-you-should-know-about/183459> accessed 19 August 2025.

[12] Anis Zalani, 'AI scams are getting real: Here are the cases happening in Malaysia that you should know about' *Malaymail* (Malaysia, 4 August 2025) <https://www.malaymail.com/news/malaysia/2025/08/04/ai-scams-are-getting-real-here-are-the-cases-happening-in-malaysia-that-you-should-know-about/183459> accessed 19 August 2025.

[13] Anis Zalani, 'AI scams are getting real: Here are the cases happening in Malaysia that you should know about' *Malaymail* (Malaysia, 4 August 2025) <https://www.malaymail.com/news/malaysia/2025/08/04/ai-scams-are-getting-real-here-are-the-cases-happening-in-malaysia-that-you-should-know-about/183459> accessed 19 August 2025.

[14] 'Viral video of border control agency chief fake' *FMT Reporters* (Malaysia, 6 August 2025) <https://www.freemalaysiatoday.com/category/nation/2025/08/06/viral-video-of-border-control-agency-chief-fake> accessed 19 August 2025.

Shockingly, in the first three months of 2025, the country recorded 12,110 online fraud cases involving scams, with total losses amounting to RM573.7 million.[15] This further shows that the government or any relevant authority has yet to take effective steps to remedy the situation since last year. Therefore, the deepfake threat is alarming not merely for its technical sophistication but also for the far-reaching societal impact it can produce.

Furthermore, deepfake has been used as a medium for cyberbullying. A recent case involved a teenager at a prestigious private secondary school who created explicit nude images of his female classmates using AI and sold each for RM2.[16] These images caused reputational damage and severe mental distress to the victims, which may never be recoverable. The long-term trauma of tech-facilitated gender-based violence, particularly against girls and women, can be worsened by the multiplier effect of digital distribution, where such images may be viewed thousands of times.

All of these real-life examples in Malaysia demonstrate that the deepfake threat is an urgent issue requiring immediate action. The financial, social, and psychological harms it has caused are spiraling out of control, and the threat will not subside unless appropriate measures are implemented to address it.

5. Evaluation on Malaysia's Current Legal Framework

Before proposing solutions for handling this issue, an evaluation of the current Malaysian framework is necessary to understand what must be amended to address and combat deepfake technology in order to better protect our citizens.

[15] Anis Zalani, 'It sounded just like my brother': How deepfake voices are fuelling money scams' *Malaymail* (Malaysia, 4 August 2025) <https://www.malaymail.com/news/malaysia/2025/08/04/it-sounded-just-like-my-brother-how-deepfake-voices-are-fuelling-money-scams/183345> accessed 19 August 2025.

[16] 'NGOs call for stronger laws against AI deepfakes after Johor case' *FMT Reporters* (Malaysia, 16 April 2025) <https://www.freemalaysiatoday.com/category/nation/2025/04/16/ngos-call-for-stronger-laws-against-ai-deepfakes-after-johor-case> accessed 19 August 2025.

Criminal Law

Although Malaysia has yet to enact specific legislation regulating AI misuse involving deepfakes,[17] victims of such cases may seek criminal redress through fragmented provisions in existing laws.

The main provision is the **Communications and Multimedia Act 1998** ("**CMA 1998**"). **Section 211(1) of the CMA 1998** prohibits the distribution of any content that is indecent, obscene, false, menacing, or offensive, with intent to annoy, abuse, threaten, or harass any person. Whoever contravenes this provision commits an offense and is punishable with a fine not exceeding RM50,000, or imprisonment for a term not exceeding one year, or both. A further fine of RM1,000 applies for every day or part of a day the offense continues after conviction.[18] This provision essentially covers the distribution of deepfake content with malicious intent that causes harm to the victim. The further fine is intended to deter repeat offenses. However, the burden of proof lies on the victim to establish malicious intent and that the content was offensive. If the distribution was reckless or negligent, and the harm was insignificant such that it does not fit within the prescribed words "annoy, abuse, threaten, or harass," the victim would not succeed under this provision. Moreover, the provision has a narrow scope as it only penalizes distribution, not creation. In other words, if one person merely creates the fake content while another distributes it, the creator may not face liability.

Despite this, **Section 233 of the CMA 1998** covers the loophole in **Section 211** by holding the creator of deepfake content liable if made with malicious intent to offend the victim, with the same punishment as **Section 211**, as stated in **Section 233(3)**. This provision is broad enough to cover developers whose tools are used to create deepfake content, even if it is not obvious that the tools are intended to deceive or serve improper purposes. However, the recent amendment striking down the words "offensive" and "annoy," though intended to align with the **Federal**

[17] Tan Zec Kie, Chong Shao Zheng, Kuek Chee Ying and Tay Eng Siang, 'Individual Legal Protection in the Deepfake Technology Era' (2023) Proceedings of the 3rd International Conference on Law and Digitalization 2023 <https://www.atlantis-press.com/article/125995062.pdf> accessed 20 July 2025.

[18] Communications and Multimedia Act 1998 (Act 588), Malaysia, s. 211(2).

Constitution,[19] may have the effect of narrowing the scope of **Section 233**, thereby making it more difficult for victims to prove bad intention in deepfake incidents.[20]

Furthermore, several provisions in the **Penal Code** are wide enough to support criminal proceedings against deepfake misuse. These include offenses of cheating, extortion, making and circulating obscene objects (including pornography), and defamation. **Section 292 of the Penal Code**, although drafted to criminalize the making, circulation, or commercial exploitation of obscene objects, also covers AI-generated deepfake pornography. This applies even if it was created without malicious intention, as the act itself is sufficient to fall under this provision unless proven otherwise to be for *bona fide* purposes. Offenders are punishable with imprisonment for a term that may extend to three years, a fine, or with both.

Section 383 of the Penal Code is relevant where perpetrators use deepfake content that could be damaging to the victim or their close ones to instill fear and induce the victim to deliver property or valuable security to prevent publication of the content. This would constitute extortion, punishable with imprisonment up to ten years, or with a fine, or with whipping, or with any two of such punishments. However, the standard of proof is higher, as it requires a ransom demand using the fake content, coupled with the victim's genuine and reasonable belief that it could cause harm, thereby inducing the victim to comply. Hence, if the victim was not in fear or the content was known to be AI-generated, this provision would not apply.

Sections 415 and 416 of the Penal Code cover scam cases where perpetrators fraudulently or dishonestly induce victims to deliver property, or impersonate someone's identity through deepfakes to deceive the victim. The punishment for cheating is imprisonment up to five years, a fine, or with both. If fake content is misused to defame someone, the perpetrator could have committed an offence under **Section 499 of the Penal Code**. The burden, however, lies on the victim to prove that the content was published by the perpetrator and that it lowered the victim's reputation, whether directly or indirectly, in the estimation of others. Punishment includes imprisonment up to two years, a fine, or with both.

[19] Communications and Multimedia (Amendment) Act 2025.

[20] 'Why the Court of Appeal Was Right to Strike Down "Annoy" and "Offensive" in Section 233' *Newswav* (Malaysia, 25 August 2025) <https://newswav.com/article/why-the-court-of-appeal-was-right-to-strike-down-annoy-and-offensive-in-sec-A2508_0ksfeH> accessed 27 August 2025.

Additional protection is afforded to children, who are particularly vulnerable. **Sections 4 to 10 of the Sexual Offences Against Children Act 2017** criminalize the making, producing, exchanging, publishing, accessing, or directing the making of child pornography.[21] **Section 4** defines "child pornography" broadly to include any representation, whether visual, audio, or written, depicting a child, whether realistic or graphic images of a child engaged in sexually explicit conduct. Thus, deepfake child pornography falls within this scope regardless of whether the child is real or synthetic, and regardless of whether the content is created, shared, or possessed. Penalties include imprisonment and whipping, which are more severe than those in the **Penal Code**, underscoring the heightened need to protect children.

Although these provisions may appear confusing and are scattered across various statutes, an independent AI statute may not be necessary, as the existing provisions are sufficiently wide to cover deepfakes misuse. A new law might only create further loopholes and confusion for the public.

Civil Law

On the other hand, victims may find it challenging to pursue civil redress, as there are limited options available to them.[22] One possible avenue is under the **Defamation Act 1957**. If deepfake content is published and made so highly realistic that it could deceive a reasonable person into believing it to be authentic, and is clearly identifiable as referring to the victim, thereby damaging the victim's reputation, the victim may take action under the Act to claim damages. However, this action can only be raised against the publisher of the fake content, not the developer or creator. The standard of proof remains relatively heavy for the victim.

Another possible action arises under the **Copyright Act 1987**, where if the fake content involves someone's work, whether literary, musical, or artistic works, films, sound recordings, or broadcasts, the owner may potentially seek action under copyright law against the infringer.

[21] Tan Zec Kie, Chong Shao Zheng, Kuek Chee Ying and Tay Eng Siang, 'Individual Legal Protection in the Deepfake Technology Era' (2023) Proceedings of the 3rd International Conference on Law and Digitalization 2023 <https://www.atlantis-press.com/article/125995062.pdf> accessed 20 July 2025.

[22] Ainin Wan Salleh, 'Give deepfake victims own legal recourse, says lawyer' *FMT Reporters* (Malaysia, 31 July 2025) <https://www.freemalaysiatoday.com/category/nation/2025/07/31/give-deepfake-victims-own-legal-recourse-says-lawyer> accessed 20 August 2025.

However, copyright law focuses on protecting originality. If the deepfake content, though created using another person's WhatsApp recordings or photographs, is sufficiently altered such that it does not resemble the original works and possess its own originality, it may not qualify as copyright infringement unless substantial similarity can be shown. Given these limitations, victims may be inclined to seek redress under the **Personal Data Protection Act 2010** ("**PDPA 2010**"), which grants individuals control over their personal data and provides recourse against misuse of personal data obtained in commercial transactions without consent.[23] The Act defines "personal data" broadly, which may include biometric identifiers such as facial images or voice recordings. Accordingly, the **PDPA 2010** indirectly grants victims to take action against those who misuse biometric data to create deepfakes, even if not done maliciously. This liability may extend not only to the misuser of the fake content but also to platform providers who collected and used personal data without consent. However, the **PDPA 2010** does not provide victims with a direct right to enforce civil actions for compensation in cases of privacy breaches.[24] Instead, remedies are limited to criminal actions, and at present, the only recourse available to an aggrieved data subject is to lodge a complaint with the Commissioner, alleging non-compliance with the **PDPA 2010**.[25]

6. Other Existing Non-Legal Loopholes

Even though there are regulations governing deepfake crimes, they may be impractical, as Malaysia still faces a major and fundamental problem in identifying the perpetrators, who are often untraceable. If the perpetrator cannot be found, victims will not be able to take any action against them. It may be suggested that the real identity of the perpetrator can be uncovered by tracing the IP address used to publish the fake content. However, what if the misuser employs a fake identity or even AI technology to conceal their tracks? For instance, in phone call scams, the deepfake

[23] Orima Melati Davey and Levin Sauerwein, 'Deepfake in Online Fraud Cases: The Haze of Artificial Intelligence's Accountability Based on the International Law' (2023) Sriwijaya Crimen and Legal Studies <https://www.researchgate.net/publication/389260078_DEEPFAKE_IN_ONLINE_FRAUD_CASES_THE_HAZE_OF_ARTIFICIAL_INTELLIGENCE'S_ACCOUNTABILITY_BASED_ON_THE_INTERNATIONAL_LAW> accessed 26 August 2025.

[24] Fadhilah Abdul Ghani, Syahirah Mohd Shabri, Maizatul Akmar Mohd Rasli, Nurulhuda Ahmad Razali and Emir Hambali Ahmad Shuffri 'An Overview of the Personal Data Protection Act 2010 (PDPA): Problems and Solutions' (2020) 12(4) Global Business and Management Research: An International Journal 559. <https://www.gbmrjournal.com/pdf/v12n4/V12N4-55.pdf> accessed 29 July 2025.

[25] Ainin Wan Salleh, 'Give deepfake victims own legal recourse, says lawyer' *FMT Reporters* (Malaysia, 31 July 2025) <https://www.freemalaysiatoday.com/category/nation/2025/07/31/give-deepfake-victims-own-legal-recourse-says-lawyer> accessed 20 August 2025.

misuser could collaborate with hackers to infiltrate a victim's family member's account, then clone the family member's voice to deceive the victim into giving money. How would the government trace the perpetrator if the number used for the deepfake is not even registered under the perpetrator's name, but instead under the victim's family member's? Hence, the critical question arises: is our system sufficiently equipped with the IT expertise to trace them?

Furthermore, even though the perpetrator is traced, what if the act is done outside of Malaysian jurisdiction? Will these provisions apply to hold the perpetrator accountable to the victim? This is a matter that should be addressed. As mentioned earlier, AI technology continues to advance rapidly, making it easier and more convenient in creating deepfake. The public's ability to detect such fabrications simply cannot keep pace. Although tools to generate deepfake are widely available, they are often not accompanied by built-in detection systems to identify their own outputs. As a result, fake content can be produced without leaving any trace of forgery.

7. Other Jurisdictions

Singapore recently passed the **Elections (Integrity of Online Advertising) (Amendment) Bill**, which seeks to combat the publication of AI-generated content that realistically depicts a political candidate as saying or doing something they did not, with the intent of influencing public voting in elections.[26] Although Indonesia has yet to implement legislation specifically targeting AI, the Minister of Communications and Informatics (MOCD) issued **Circular Letter No. 9 of 2023 on AI Ethics**,[27] which provides ethical guidelines for companies in formulating internal AI policies, particularly regarding data use and accountability. It also governs AI-based activities to hold AI operators responsible for misconduct. Additionally, the Indonesian Financial Services Authority (OJK) released a code of ethics for responsible and trustworthy AI in the financial technology

[26] Chin Soo Fang, 'Bill passed to counter digitally manipulated content, deepfakes during elections' *The Straits Times* (Malaysia, 15 October 2024) <https://www.straitstimes.com/singapore/politics/bill-passed-to-counter-digitally-manipulated-content-deepfakes-during-elections> accessed 24 August 2025 .

[27] Ayik Candrawulan Gunadi, Mahiswara Timur and Marsher Miyata, 'Legal 500 Country Comparative Guides 2025' (ABNR Counsellors at Law, 2025) <https://www.legal500.com/guides/chapter/indonesia-artificial-intelligence/?export-pdf#:~:text=The%20Indonesian%20Financial%20Services%20Authority,AI%20by%20financial%20technology%20providers> accessed 26 August 2025.

industry, setting out key principles for fintech providers.[28] These measures ensure that AI is used ethically and mandate AI-generated content to be clearly disclosed as such.

On the other hand, Taiwan has taken early steps to address deepfake issues through the **Fraud Crime Prevention Act**, which aims to prevent fraud by requiring industries such as banks, e-commerce platforms, and telecom companies to comply with strict rules.[29] This source-based fraud prevention mechanism obliges businesses to block fraud at its origin rather than only targeting fraudsters. Obligations include adopting fraud prevention measures, verifying customer identities, promoting fraud awareness, and cooperating with law enforcement.

South Korea faces serious issues with deepfake pornography. In 2019, a notorious sex ring was uncovered that used deepfakes to exploit dozens of victims, including minors. The ringleader, Cho Ju-bin, was sentenced to 42 years in prison.[30] This illustrates that South Korea had already introduced laws criminalizing the creation and distribution of sexually explicit deepfake content without consent. In 2024, the country further criminalized the consumption of deepfake pornography to curb its spread. Ongoing reforms have been proposed to the **Act on Special Cases Concerning the Punishment of Sexual Crimes**, including raising the maximum sentence for producing and distributing deepfake pornographic material from five to seven years, as well as clarifying the government's responsibility to remove explicit content and support victims.[31]

The United States recently enacted the **Take It Down Act**, which makes it illegal to share non-consensual explicit images online, whether real or AI-generated. The law also obliges tech platforms to remove such images within 48 hours of receiving a complaint.[32] Major platforms such

---

[28] Ayik Candrawulan Gunadi, Mahiswara Timur and Marsher Miyata, 'Legal 500 Country Comparative Guides 2025' (ABNR Counsellors at Law, 2025) <https://www.legal500.com/guides/chapter/indonesia-artificial-intelligence/?export-pdf#:~:text=The%20Indonesian%20Financial%20Services%20Authority,AI%20by%20financial%20technology%20providers> accessed 26 August 2025.

[29] 'Navigating Taiwan's New Anti-Fraud Law: How Can Businesses Comply?' (*Tookitaki*, 13 December 2024) <https://www.tookitaki.com/blog/navigating-taiwans-new-anti-fraud-law-how-can-businesses-comply> accessed 25 July 2025.

[30] R. Loheswar, 'War on deepfakes: What Malaysia can learn from the rest of the world' *Malaymail* (Malaysia, 5 November 2024) <https://www.malaymail.com/news/malaysia/2024/11/05/war-on-deepfakes-what-malaysia-can-learn-from-the-rest-of-the-world/154811> accessed 3 August 2025.

[31] Yoonjung Seo and Mike Valerio, 'Deepfake porn is destroying real lives in South Korea' *CNN* (Seoul, South Korea, 29 April 2025) <https://edition.cnn.com/2025/04/25/asia/south-korea-deepfake-crimes-intl-hnk-dst> accessed 12 July 2025.

[32] Clare Duffy, 'Victims of explicit deepfakes will now be able to take legal action against people who create them' *CNN* (New York, 19 May 2025) <https://edition.cnn.com/2025/05/19/tech/ai-explicit-deepfakes-trump-sign-take-it-down-act#:~:text=Victims%20of%20explicit%20deepfakes%20will,against%20people%20who%20create%20them&text=After%20months%20of%20outcry%20over,will%20make%20sharing%20them%20illegal.&text=In%20recent%20years%2C%20people%20ranging%20from%20Taylor%20Swift%20and%20Rep> accessed 12 July 2025.

as Google, Meta, and Snapchat already provide mechanisms for users to request the removal of explicit images. Some have also partnered with non-profit organizations such as StopNCII.org and Take It Down, which facilitate cross-platform removal of such images, though not all sites cooperate. Apple and Google have also removed AI services that transform clothed images into nude manipulated ones from their app stores and search results.[33]

Denmark is in the process of developing new deepfake legislation as part of its digital copyright law, granting individuals legal rights over their biometric data, including body, facial features, and voice, to suppress the creation and dissemination of unauthorized deepfakes.[34] Rather than solely focusing on preventing harmful content, the proposed bill empowers individuals to take action against perpetrators and online platforms that misuse their biometric data without consent.[35] In Malaysia, however, biometric data falls within the scope of the **PDPA 2010**, not the **Copyright Act 1987**. Accordingly, amendments to the **PDPA 2010** could be considered to avoid confusion, as the Act already governs personal data.

From the above, it is evident that several jurisdictions have encountered deepfake threats earlier than anticipated and have introduced AI-specific regulations to address them. By contrast, Malaysia risks appearing outdated for not amending its statutes accordingly. Therefore, policymakers should look to these examples when considering reforms to effectively tackle AI-related threats.

[33] Clare Duffy, 'Victims of explicit deepfakes will now be able to take legal action against people who create them' *CNN* (New York, 19 May 2025) <https://edition.cnn.com/2025/05/19/tech/ai-explicit-deepfakes-trump-sign-take-it-down-act#:~:text=Victims%20of%20explicit%20deepfakes%20will,against%20people%20who%20create%20them&text=After%20months%20of%20outcry%20over,will%20make%20sharing%20them%20illegal.&text=In%20recent%20years%2C%20people%20ranging%20from%20Taylor%20Swift%20and%20Rep> accessed 12 July 2025.

[34] Ainin Wan Salleh, 'Give deepfake victims own legal recourse, says lawyer' *FMT Reporters* (Malaysia, 31 July 2025) <https://www.freemalaysiatoday.com/category/nation/2025/07/31/give-deepfake-victims-own-legal-recourse-says-lawyer> accessed 20 August 2025.

[35] Andrea Willige, 'Deepfake legislation: Denmark moves to protect digital identity' (*World Economic Forum*, 30 July 2025) <https://www.weforum.org/stories/2025/07/deepfake-legislation-denmark-digital-id/#:~:text=In%20Denmark%2C%20the%20government%20is,including%20their%20appearance%20and%20voice> accessed 22 August 2025.

8. Recommendations

Denmark's proposed bill has given us an idea to reform our **PDPA 2010** to explicitly provide everyone with legal rights and control over their personal data,[36] including appearance and voice.[37] The **PDPA 2010** was not designed with AI-generated threats in mind and therefore offers limited protection to individuals whose image, voice, and identity are manipulated without consent, especially when the Act relies heavily on the Commissioner to take action instead of the victims themselves. Therefore, instead of creating new laws, the existing laws, particularly the **PDPA 2010**, should be strengthened to state that biometric and synthetic data such as facial scans and voice prints used in AI-generated content fall under the Act's purview, and a specific provision should be imposed to allow private actions against misusers of other's biometric data without consent. This would sufficiently grant anyone victimized by these AI manipulations the right to take swift legal action against the perpetrators and even the platforms that publish such content, ensuring that it is promptly removed and that those responsible are penalized. This aims not only to avoid the malicious use of synthetic media by others but also to ensure that such use must be consented to by the data owner. Malaysia currently lacks civil redress for victims against deepfake perpetrators; this amendment would fill that gap.[38]

The recent new amendment to the **CMA 1998** introduced a new provision namely **Section 236A**, which grants the right of private action to anyone to seek civil relief against alleged offenders charged with an offence under **Sections 235 and 236**.[39] This revision has invited civil redress to the **CMA 1998**, giving hope to deepfake victims to expect the same right to be accorded to them against deepfake perpetrators found guilty under **Sections 211 and 233**. This says that if

---

[36] Tan Zec Kie, Chong Shao Zheng, Kuek Chee Ying and Tay Eng Siang, 'Individual Legal Protection in the Deepfake Technology Era' (2023) Proceedings of the 3rd International Conference on Law and Digitalization 2023 <https://www.atlantis-press.com/article/125995062.pdf> accessed 20 July 2025.

[37] Ainin Wan Salleh, 'Give deepfake victims own legal recourse, says lawyer' *FMT Reporters* (Malaysia, 31 July 2025) <https://www.freemalaysiatoday.com/category/nation/2025/07/31/give-deepfake-victims-own-legal-recourse-says-lawyer> accessed 20 August 2025.

[38] Fadhilah Abdul Ghani, Syahirah Mohd Shabri, Maizatul Akmar Mohd Rasli, Nurulhuda Ahmad Razali and Emir Hambali Ahmad Shuffri 'An Overview of the Personal Data Protection Act 2010 (PDPA): Problems and Solutions' (2020) 12(4) Global Business and Management Research: An International Journal 559 <https://www.gbmrjournal.com/pdf/v12n4/V12N4-55.pdf> accessed 29 July 2025.

[39] Ariff Firman Bin Mohd Sidek and Nur Aliya Syaffa Binti Johari, 'Key Updates to the Communications and Multimedia (Amendment) Act 2025 and Recommendations for Industry Preparedness' (2025) LAW Partnership, LAW Updates <https://law-partnership.com/wp-content/uploads/public_files/Key%20Updates%20to%20the%20CMA%202025.pdf> accessed 31 August 2025.

proven guilty under **Sections 211 or 233**, not only criminal sanctions will be imposed on the perpetrators, but the victims could take actions against the perpetrators to claim civil remedies such as damages and injunction.

Furthermore, in our evaluation of the current legal framework, our comment is that it focuses on the reactive aspect instead of the proactive. Proactive measures should be taken, as they could stop deepfakes at the root instead of addressing them only after harm has occurred. Indonesia would be a good example with its proactive measure, which is to focus on areas related to accountability and the integrity of AI developers and users to encourage ethical use of AI. Malaysia could model its approach by strengthening the **ASEAN Guide on AI Governance and Ethics** in order to regulate the use of AI.[40] This would ensure that AI is kept under control, similar to over-the-top platforms.[41] All companies or individuals providing AI services capable of generating deepfakes must be legally registered and identified. This is to hold them accountable in the event of any deepfake cases. In other words, what our country is currently facing is that these platforms lack restrictions, which allows unauthorized use of personal data and identity exploitation.

Strict rules and regulations should also be imposed on these AI service providers, including social media platforms, in allowing their services.[42] For example, detection tools should be provided along with their services. All AI-generated content should be watermarked as AI-generated, with a deepfake disclaimer displayed before posting, and these AI detection tools should be readily available to the general public.[43] Moreover, the government should invest in IT companies that could offer detection tools as applications easily accessible to the public,[44] such as

[40] 'NGOs call for stronger laws against AI deepfakes after Johor case' *FMT Reporters* (Malaysia, 16 April 2025) <https://www.freemalaysiatoday.com/category/nation/2025/04/16/ngos-call-for-stronger-laws-against-ai-deepfakes-after-johor-case> accessed 19 August 2025.

[41] 'Tackle deepfake abuse with legal framework, say experts' *The Star* (Malaysia, 20 August 2025) <https://www.thestar.com.my/news/nation/2024/08/20/tackle-deepfake-abuse-with-legal-framework-say-experts> accessed 19 August 2025.

[42] 'Tackle deepfake abuse with legal framework, say experts' *The Star* (Malaysia, 20 August 2025) <https://www.thestar.com.my/news/nation/2024/08/20/tackle-deepfake-abuse-with-legal-framework-say-experts> accessed 19 August 2025.

[43] Law Kian Seng, Normaisharah Mamat, Hafiza Abas and Wan Noor Hamiza Wan Ali, 'AI Integrity Solutions for Deepfake Identification and Prevention' (2024) 12(1) Open International Journal of Informatics (OIJI) 35–46 <https://oiji.utm.my/index.php/oiji/article/view/297/207> accessed 12 July 2025.

[44] Tan Zec Kie, Chong Shao Zheng, Kuek Chee Ying and Tay Eng Siang, 'Individual Legal Protection in the Deepfake Technology Era' (2023) Proceedings of the 3rd International Conference on Law and Digitalization 2023 <https://www.atlantis-press.com/article/125995062.pdf> accessed 20 July 2025.

downloadable apps that function like virus detectors, giving out warnings on smartphones when they detect potential deepfake content.

Moreover, platforms, including banks and payment platforms, could implement a cooling-off period of 24 or 48 hours to avoid deepfake scams.[45] This means that when a potential victim receives digital content that either asks for urgent money or lures them into clicking on it to transfer money, the bank or payment platform should not allow an immediate transfer of funds, especially when it involves large amounts. This would allow the potential victim extra time to review, verify red flags, consult the platform, or even cancel if the content appears suspicious during the cooling-off period. Otherwise, the transfer could become irreversible.

By placing accountability on platforms, they would be subject to a form of digital insurance against damages caused by deepfake-related cases. In other words, if victims suffer monetary or reputational damages, they would be able to hold the platforms accountable for their losses by claiming damages. This would be an effective method to prevent deepfakes by creating incentives for platforms to adopt robust verification mechanisms, monitoring tools, and fraud-prevention measures, thereby shifting liability to the deepfake misusers.

Undeniably, education and social media literacy among the public would be the best defense against deepfakes. The public, especially the younger generation who are more familiar with technology, should be educated rather than prohibited from using AI.[46] They should know how AI works and how and why deepfakes are created. With this knowledge, the public would remain vigilant, especially when there are available supporting resources such as official channels or trusted sources, allowing them to independently verify digital content if it seems suspicious.[47] They should approach online content with a healthy dose of scepticism, particularly when it involves endorsements from celebrities and public figures.

---

[45] 'Tackle deepfake abuse with legal framework, say experts' *The Star* (Malaysia, 20 August 2025) <https://www.thestar.com.my/news/nation/2024/08/20/tackle-deepfake-abuse-with-legal-framework-say-experts> accessed 19 August 2025.

[46] Anis Zalani, 'That voice isn't real: Why media literacy matters in the age of deepfake and AI scams' *Malaymail* (Malaysia, 5 August 2025) <https://www.malaymail.com/news/malaysia/2025/08/05/that-voice-isnt-real-why-media-literacy-matters-in-the-age-of-deepfake-and-ai-scams/183565> accessed 19 August 2025.

[47] Anis Zalani, 'It sounded just like my brother': How deepfake voices are fuelling money scams' *Malaymail* (Malaysia, 4 August 2025) <https://www.malaymail.com/news/malaysia/2025/08/04/it-sounded-just-like-my-brother-how-deepfake-voices-are-fuelling-money-scams/183345> accessed 19 August 2025.

In the case of AI-generated voice cloning, scammers often impersonate family members, friends, or colleagues to create a false sense of urgency, pushing victims to make quick decisions involving money. However, the public should understand that there is no such urgency that does not allow some time to verify the reliability of the content. Instead of being pressured by the urgency, potential victims should verify by other means, such as video calling to confirm the caller's identity, asking to meet in person to hand over money instead of transferring online, or creating a safe code among family members.

9. Conclusion

The above discussions briefly described an overview of the current legal framework in Malaysia in addressing deepfake threats. The great potential danger that deepfakes can bring necessitates Malaysia to move forward instead of relying solely on existing regulations to govern them. Our current legal framework focuses mainly on reactive measures, which lack effectiveness in stopping deepfakes at their roots, especially given the difficulties in identifying the true identity of deepfake creators. Preventive measures should be introduced together with civil remedies to compensate victims for their losses, such as injunctions and damages.