Justice Administration of Tehran Province

Judgment

https://www.Ghazavat.org

Online ISSN: 3060-7809
Print ISSN: 1735-1227
Volum:26 – Issue:126
Summer 2026

Legal Challenges in Judicial Proceedings for Addressing Crimes Committed Through Deepfakes

Mona Hosseinpouri[1]

1. Assistant Professor, Department of Law, Kermanshah Branch, Islamic Azad University, Kermanshah, Iran. hosseinpouri.m.1366@iau.ac.ir

Article Info

Article type:
Scientific Article

Date Received:
2025/04/29
Date Revised:
2025/12/30
Date Accepted:
2026/05/02

Keywords:
Crime Commission; Legal Challenges; Deepfakes; Judicial Proceedings

Abstract

Deepfake refers to technologies that, employing deep learning and neural networks to manipulate audio, images or video content in ways that make it appear authentic, while in fact being entirely fabricated. This study, employing a descriptive-analytical methodology, examines the legal challenges of judicial proceedings in cases arising from the commission of crimes through the use of deepfakes. The necessity of this research stems from the unprecedented challenges faced by judicial systems in determining criminal liability, establishing the mens rea element, assessing the evidentiary validity of digital content where traditional approaches in criminal law are insufficient to address these complexities. Findings indicate that confronting these challenges requires the development of novel legal-technical frameworks. Such frameworks should include drafting specialized laws for deepfake-related crimes, applying strict liability for producers, enhancing transparency of system operations through techniques such as digital watermarking, and providing enhanced judicial training and expertise in the field of deepfakes. The present study emphasizes that an intelligent integration of legal principles, advanced monitoring technologies and ethical frameworks can contribute to establishing a safer and more equitable digital ecosystem. This integrated approach not only protects individual rights but also provides the necessary space for the growth of transformative technologies and can serve as a model for addressing future challenges posed by emerging technologies.

How To Cite

Hosseinpouri, Mona. (2026). Legal Challenges in Judicial Proceedings for Addressing Crimes Committed Through Deepfakes. Journal of Judgment, 126(2), 21-34.
http//doi.org/ 10.22034/judg.2026.2059297.1448

DOI

10.22034/judg.2026.2059297.1448

Publisher

Publications of the Judiciary of Tehran Province

آموزش دادگستری
استان تهران

فصلنامه قضاوت

https://www.Ghazavat.org

شاپای الکترونیکی: ۷۸۰۹-۳۰۶۰
شاپا چاپی: ۱۲۲۷-۱۷۳۵
دوره: ۲۶ ـ شماره: ۱۲۶
تابستان ۱۴۰۵

چالش‌های حقوقی رسیدگی قضایی در بررسی ارتکاب جرایم به کمک جعل عمیق

منا حسین‌پوری[۱]

۱. استادیار گروه حقوق، واحد کرمانشاه، دانشگاه ازاد اسلامی، کرمانشاه، ایران. رایانامه: hosseinpouri.m.1366@iau.ac.ir

اطلاعات مقاله

نوع مقاله: علمی

تاریخ دریافت:
۱۴۰۴/۰۲/۰۹
تاریخ بازنگری:
۱۴۰۴/۱۰/۰۹
تاریخ پذیرش:
۱۴۰۵/۰۲/۱۲

کلیدواژه:
ارتکاب جرم، چالش‌های حقوقی، جعل عمیق، دادرسی.

چکیده

جعل عمیق به فناوری‌هایی اطلاق می‌شود که با استفاده از یادگیری عمیق و شبکه‌های عصبی، محتوای صوتی، تصویری یا ویدئویی را به گونه‌ای دستکاری می‌کند که واقعی به نظر برسد، اما کاملاً جعلی است. در این پژوهش با روش توصیفی ـ تحلیلی، چالش‌های حقوقی رسیدگی قضایی به پرونده‌های ناشی از ارتکاب جرم به کمک جعل عمیق (دیپ‌فیک) بررسی می‌شود. ضرورت انجام این تحقیق آن است که نظام‌های قضایی با چالش‌های بی‌سابقه‌ای در تعیین مسئولیت کیفری، اثبات عنصر ذهنی و اعتبارسنجی ادله مواجه شده‌اند و رویکردهای سنتی حقوق کیفری پاسخ‌گوی این پیچیدگی‌ها نیستند. بر اساس یافته‌ها، برای مواجهه با این چالش‌ها، گسترش چهارچوب‌های نوین حقوقی ـ فنی ضروری است. این چهارچوب‌ها باید شامل تدوین قوانین اختصاصی برای جرایم دیپ‌فیک، اعمال نظریهٔ مسئولیت محض برای تولیدکنندگان، شفاف‌سازی عملکرد سیستم‌ها از طریق روش‌هایی مانند واترمارکینگ، و آموزش و افزایش تخصص قضایی در زمینهٔ دیپ‌فیک باشند. در پژوهش حاضر تأکید می‌شود که ترکیب هوشمندانهٔ اصول حقوقی، فناوری‌های پیشرفتهٔ نظارتی و چهارچوب‌های اخلاقی می‌تواند به برقراری اکوسیستم دیجیتالی امن‌تر و عادلانه‌تری بینجامد. این رویکرد یکپارچه علاوه بر آنکه از حقوق افراد محافظت می‌کند، فضای لازم برای رشد فناوری‌های تحول‌آفرین را نیز فراهم می‌سازد و می‌تواند به‌مثابهٔ الگویی برای مواجهه با چالش‌های آیندهٔ فناوری‌های نوین استفاده شود.

استناد
حسین‌پوری، منا. (۱۴۰۵). چالش‌های حقوقی رسیدگی قضایی در بررسی ارتکاب جرایم به کمک جعل عمیق، فصلنامه قضاوت، ۱۲۶(۲)، ۲۱-۳۴.
http//doi.org/ 10.22034/judg.2026.2059297.1448

DOI 10.22034/judg.2026.2059297.1448

ناشر انتشارات دادگستری کل استان تهران

مقدمه

پیشرفت شتابان هوش مصنوعی در دهه‌های اخیر، به‌ویژه با اتکا بر الگوریتم‌های یادگیری عمیق، شبکه‌های مولد تخاصمی[۱] و مدل‌های زبانی و چندوجهی بزرگ،[۲] به ظهور فناوری‌هایی منجر شده است که توانایی تولید محتوای صوتی، تصویری و ویدئویی با سطحی بی‌سابقه از واقع‌نمایی را دارند؛ به‌گونه‌ای که در بسیاری از موارد، تشخیص این محتواها از داده‌های واقعی برای انسان و حتی سامانه‌های فنی بسیار دشوار یا تقریباً ناممکن شده است. این فناوری که در ادبیات علمی با عنوان «جعل عمیق» یا «دیپ‌فیک» شناخته می‌شود، از منظر فنی مبتنی بر معماری‌های عصبی عمیقی[۳] است که قابلیت مدل‌سازی توزیع‌های احتمالی پیچیده را در داده‌های ورودی دارند و با حفظ ویژگی‌های آماری داده‌های آموزشی، نمونه‌های جدید و به ظاهر معتبر تولید می‌کنند (Binns and Veale, 2021: 3).

هرچند این فناوری در برخی حوزه‌ها مانند آموزش، پزشکی و سرگرمی کاربردهای مشروع و مفیدی دارد، ظرفیت بالای آن برای سوءاستفاده، به‌ویژه در بستر فضای سایبری، موجب شده است که دیپ‌فیک به یکی از مهم‌ترین تهدیدات نوظهور علیه امنیت فردی، اجتماعی و نهادی تبدیل شود. بر اساس مطالعات تجربی، استفاده از محتوای تولیدشده از طریق هوش مصنوعی، به دلیل شخصی‌سازی پیشرفته و واقع‌نمایی بالا، می‌تواند نرخ موفقیت حملات فیشینگ و کلاهبرداری را تا حدود ۳۰درصد افزایش دهد و بدین ترتیب، سازوکارهای تشخیص سنتی را که بر الگوهای ثابت یا نشانه‌های ظاهری متکی هستند، با چالش جدی مواجه سازد (Helm et al., 2020: 70).

نمودهای عینی این تهدیدات شامل استفاده از صداهای جعلی برای دورزدن سامانه‌های احراز هویت بیومتریک، تولید ویدئوهای دیپ‌فیک برای جعل هویت مدیران ارشد سازمان‌ها، و ساخت اسناد و مدارک جعلی با ظاهری کاملاً حرفه‌ای است که می‌تواند در فرایندهای حقوقی و اداری مورد استناد قرار گیرد. طبق تحقیقات، این شیوه‌ها به‌ویژه در حملات هدفمند علیه سازمان‌های مالی و زیرساخت‌های حیاتی که در آن‌ها عامل انسانی ضعیف‌ترین حلقهٔ امنیتی محسوب می‌شود، اثربخشی بالایی دارند (Walters and Novak, 2021: 124) و از منظر علوم شناختی، موفقیت آن‌ها ناشی از بهره‌برداری هوشمندانه از سوگیری‌های شناختی و محدودیت‌های پردازشی نظام ادراکی انسان است. در چنین شرایطی، رشد سریع فناوری دیپ‌فیک با توانایی تولید محتوایی که تشخیص

---

۱. معماری یادگیری عمیق متشکل از دو شبکهٔ رقیب که یکی دادهٔ جعلی تولید می‌کند و دیگری اصالت آن را تشخیص می‌دهد تا در نهایت محتوایی بسیار واقع‌نما ایجاد شود.

۲. مدل‌های هوش مصنوعی مبتنی بر داده‌های بزرگ که قادر به درک و تولید زبان طبیعی با سطحی نزدیک به انسان هستند.

۳. ساختارهای محاسباتی مبتنی بر لایه‌های متعدد نورونی که برای مدل‌سازی روابط غیرخطی پیچیده به کار می‌روند.

آن از واقعیت تقریباً غیرممکن شـده اسـت، چالش‌های بی‌سابقه‌ای را به‌ویژه در مرحلهٔ دادرسی و رسیدگی قضایی پدید آورده است (Nayak et al., 2024: 96).

در چهارچوب حقوق کیفری کلاسـیک، مسـئولیت کیفری مبتنی بر عناصری مانند تقصیر، قصد مجرمانه و رابطهٔ سـببیت اسـت، حال آنکه در جرایم مبتنی بر دیپ‌فیک، به دلیل ماهیت غیرمستقیم، توزیع‌شـده و فناورانهٔ فرایند ارتکاب جرم، احراز این عناصر با ابهام‌های جدی همراه اسـت. برای مثال، در پرونده‌هایی که ویدئویی دیپ‌فیک با هدف تخریب شـخصیت مورد اسـتفاده قرار می‌گیرد، این پرسش اساسی مطرح می‌شود که مسئولیت کیفری متوجه چه کسی است؛ توسعه‌دهندهٔ الگوریتم، کاربری که محتوا را تولید یا منتشر کرده، یا پلتفرمی که آن را میزبانی کرده است (ابراهیمی، ۱۴۰۲: ۲۷)؟

پیچیدگـی این وضعیت زمانی افزایش می‌یابد که نقش فناوری‌های تشـخیص دیپ‌فیک نیز به دلیل ماهیت «جعبهٔ سیاه» بسیاری از الگوریتم‌های یادگیری عمیق، با مسئلهٔ شفافیت و تفسیرپذیری مواجه می‌شـود؛ به‌گونه‌ای که حتی طراحان این سامانه‌ها نیز قادر به توضیح دقیق فرایند استدلالی سیسـتم در تشـخیص جعل نیسـتند (Marr and Ward, 2019: 108). این نبود شـفافیت، در بستر دادرسی عادلانه که مستلزم امکان ارزیابی، اعتراض و توجیه مستند احکام است، می‌تواند به مانعی جدی برای پذیرش ادلهٔ مبتنی بر دیپ‌فیک تبدیل شود.

از سوی دیگر، دیپ‌فیک‌ها علاوه بر آنکه در جرایم مالی مانند کلاهبرداری‌های پیشرفته، جعل اسـناد و حملات موسـوم به «ایمیل تجاری» که سـالانه میلیاردها دلار خسارت به اقتصاد جهانی وارد می‌کننـد، تأثیرگذارنـد (Gong and Li, 2024: 587)، در حوزهٔ جرایم غیرمالی نیز مانند هتک حیثیت، جعل هویت، انتساب اظهارات یا رفتارهای مجرمانه به اشخاص حقیقی و حقوقی، انتشار محتوای غیرقانونی و دسـتکاری افکار عمومی در فرایندهای دموکراتیک، پیامدهای جدی به دنبال دارند (Chac, 2020: 23; Wachter, 2024: 693).

از منظر فنی نیز، چالش تشـخیص دیپ‌فیک‌ها با ظهور نسل جدید مدل‌های مبتنی بر دیفیوژن[1] تشـدید شده است؛ به‌طوری که حتی پیشرفته‌ترین سامانه‌های تشخیصی مبتنی بر یادگیری ماشین، در بهترین حالت به دقتی حدود ۸۵ تا ۹۰درصد دست می‌یابند، که برای کاربردهای امنیتی و قضایی مهم کافی نیست (Ahmed et al., 2022: 2–3; Nayak et al., 2024: 93–94). بر اساس مجموع این تحولات، نظام حقوقی موجود، چه در سـطح قوانین ماهوی و چه در حوزهٔ آیین دادرسی، کمبودها و نارسایی‌های جدی در مواجهه با جرایم مبتنی بر جعل عمیق دارد.

بنابراین، بیان مسـئلهٔ اصلی این پژوهش آن اسـت که چگونه می‌تـوان در چهارچوب حقوق کیفـری، با حفظ اصول بنیادینی همچون قانونی‌بودن جرم و مجازات، دادرسـی عادلانه و حمایت

۱. الگوریتم‌های مولد پیشرفته‌ای که با حذف تدریجی نویز از دادهٔ تصادفی، محتوای بسیار واقعی تولید می‌کنند.

از حقوق طرفین، رسیدگی مؤثر و منصفانه در خصوص جرایمی داشت که با استفاده از فناوری دیپ‌فیک ارتکاب می‌یابند. اهمیت و ضرورت این تحقیق هم از آنجا ناشی می‌شود که بی‌توجهی به این چالش‌ها می‌تواند به تضعیف اعتماد عمومی به نظام عدالت کیفری، افزایش مصونیت مرتکبان جرایم فناورانه و گسترش بی‌ثباتی اجتماعی منجر شود (فرزین و سمیعی، ۱۴۰۲: ۲۱). در این نوشتار با بهره‌گیری از روش توصیفی ـ تحلیلی و با اتکا به منابع کتابخانه‌ای و پژوهش‌های بین‌رشته‌ای، تلاش شده است ضمن تبیین ابعاد فنی و حقوقی مسئله، ضرورت بازاندیشی در مقررات موجود و تقویت آموزش و تخصص‌گرایی در بدنهٔ قضایی برجسته شود تا زمینه برای مواجهه‌ای کارآمدتر با جرایم مبتنی بر دیپ‌فیک فراهم آید.

پیشینه و نوآوری پژوهش

۱. در پژوهش فلاح تفتی و عبدخدایی، با عنوان «تنظیم‌گری دیپ‌فیک در پرتو فقه حکومتی: الگوی فقهی ـ حقوقی برای حکمرانی بر این فناوری»، با رویکردی بین‌رشته‌ای، امکان تنظیم‌گری فناوری دیپ‌فیک بر مبنای فقه حکومتی تبیین شده است. نویسندگان ضمن تحلیل اصول فقهی مرتبط با مصلحت عمومی، نظم اجتماعی و دفع مفاسد، الگویی فقهی و حقوقی برای حکمرانی بر دیپ‌فیک ارائه و نشان داده‌اند که فقه اسلامی ظرفیت پاسخ‌گویی به چالش‌های نوپدید فناوری‌های هوش مصنوعی را دارد.

۲. رضایی و امینی در پژوهشی با عنوان «مسئولیت کیفری در دیپ‌فیک و جعل هویت تصویری»، با تمرکز بر مفهوم مسئولیت کیفری در جرایم ناشی از دیپ‌فیک و جعل هویت تصویری، ارکان قانونی، مادی و معنوی جرم را در فضای سایبری بررسی کرده‌اند. نویسندگان با تحلیل حقوق کیفری کلاسیک و تطبیق آن با ویژگی‌های فنی دیپ‌فیک، چالش‌های انتساب رفتار مجرمانه، قصد مجرمانه و رابطهٔ سببیت را واکاوی کرده و خلأهای تقنینی موجود در نظام حقوقی ایران را برجسته ساخته‌اند.

۳. فاخری و همکاران در پژوهش «مدل‌ها و روش‌های تشخیص دیپ‌فیک در هوش مصنوعی و تأثیر این پدیده بر فرهنگ اجتماعی از منظر رسانه»، با رویکردی فناورانه و اجتماعی، مدل‌ها و الگوریتم‌های تشخیص دیپ‌فیک را در هوش مصنوعی بررسی و پیامدهای فرهنگی و رسانه‌ای این پدیده را تحلیل کرده‌اند. این نویسندگان ضمن معرفی روش‌های نوین تشخیص جعل عمیق، تأثیر گسترش دیپ‌فیک در اعتماد اجتماعی، ادراک عمومی از حقیقت و نقش رسانه‌ها در بازتولید یا کنترل این پدیده را تبیین کرده‌اند.

۴. علی‌یـن در پژوهش «جرم‌انگاری دیپ‌فیک‌ها از منظر تعهدات حقوق بشری دولت‌ها»،

ضـرورت و حدود جرم‌انـگاری دیپ‌فیک‌ها را در پرتو تعهدات حقوق بشـری دولت‌ها بررسـی کرده اسـت. او با تحلیل اسناد بین‌المللی حقوق بشـر، تعارض میان آزادی بیان و حمایـت از حیثیت، حریـم خصوصی و امنیت عمومی را واکاوی کـرده و بر لزوم تدوین سیاسـت کیفری متوازن تأکید کرده است، که هم از سوءاستفاده از دیپ‌فیک جلوگیری کند و هم ناقض حقوق بنیادین افراد نباشد.

نـوآوری اصلی پژوهـش حاضر با عنوان «چالش‌های حقوقی رسـیدگی قضایی در بررسـی ارتـکاب جرایـم به کمک جعل عمیق» در تمرکز نظام‌مند و تحلیلی آن بر مرحلهٔ رسـیدگی قضایی و دادرسـی جرایم مبتنی بر دیپ‌فیک نهفته اسـت. این مرحله به‌رغم اهمیت بنیادین آن در تحقق عدالت کیفری، در ادبیات حقوقی داخلی کمتر مورد توجه مسـتقل قرار گرفته اسـت. طبق بررسی پیشینهٔ پژوهش، مطالعات انجام‌شده تاکنون عمدتاً به ابعاد ماهوی پدیدهٔ دیپ‌فیک معطوف بوده‌اند؛ به‌گونه‌ای که در برخی پژوهش‌ها بر مسئولیت کیفری و عناصر جرم در جعل هویت تصویری تمرکز شـده و چالش‌های انتسـاب رفتار مجرمانه، قصد مجرمانه و رابطهٔ سـببیت در چهارچوب حقوق کیفری کلاسـیک واکاوی شده‌اند، بدون آنکه پیامدهای این ابهام‌ها در فرایند اثبات، ارزیابی ادله و صدور رأی در مرحلهٔ دادرسی به‌طور مستقل تحلیل شوند.

در مقابل، در بخشـی دیگـر از مطالعات، به تنظیم‌گری تقنینی و فقهی دیپ‌فیک توجه شـده و تلاش شـده اسـت الگوهایی برای حکمرانی بر این فناوری ارائه شـود یا از منظر تعهدات حقوق بشری دولت‌ها، ضرورت و حدود جرم‌انگاری دیپ‌فیک‌ها بررسی شوند. همچنین برخی تحقیقات با رویکرد فناورانه یا فرهنگی، بر روش‌های تشخیص فنی دیپ‌فیک و آثار اجتماعی و رسانه‌ای آن متمرکز شـده‌اند. با وجود ارزش علمی این آثار، نقطهٔ مشـترک آن‌ها غفلت نسبی از تحلیل دقیق و مرحله‌مند چالش‌های حقوقی رسـیدگی قضایی است؛ به این معنا که مسائل کلیدی همچون اعتبار و قابلیت اسـتناد ادلهٔ دیجیتال مبتنی بر هوش مصنوعی، نقش کارشناسـی فنی در دادرسـی، حدود اتکای قاضی به سامانه‌های تشخیص دیپ‌فیک، و پیامدهای این وضعیت بر اصول بنیادین دادرسی منصفانه، به‌صورت منسجم و متمرکز بررسی قرار نشده‌اند. نوآوری پژوهش حاضر دقیقاً در پرکردن همین خلأ علمی تبلور می‌یابد؛ بدین معنا که در این پژوهش با عبور از سطح صرفاً ماهوی مسئولیت کیفری، به‌طور عمیق و بین‌رشـته‌ای چالش‌هایی واکاوی می‌شـود که فنـاوری دیپ‌فیک در فرایند رسیدگی قضایی ایجاد می‌کند و نشان می‌دهد که حتی در فرض تحقق ارکان جرم و امکان انتساب مسـئولیت کیفری، نظام دادرسی با موانع جدی در احراز اصالت ادله، ارزیابی فنی شواهد، تضمین حقوق دفاعی متهم و حفظ بی‌طرفی و شفافیت در تصمیم قضایی مواجه است.

۱. چالش‌های حقوقی در مرحلهٔ رسیدگی

در این بخش چالش‌های حقوقی که ارتکاب جرم از طریق دیپ‌فیک برای سیستم قضایی در مرحلهٔ رسیدگی ایجاد می‌کند، بررسی می‌شود.

۱-۱. ماهیت جرم‌های مبتنی بر جعل عمیق

جرایم مبتنی بر دیپ‌فیک از منظر حقوقی و جرم‌شناسی پدیده‌ای منحصربه‌فرد محسوب می‌شوند که در آن فناوری‌های پیشرفتهٔ یادگیری عمیق به ابزاری برای ارتکاب جرایم سایبری پیچیده تبدیل شده‌اند. ماهیت چندلایه و توزیع‌یافتهٔ این فناوری‌ها چالش‌های حقوقی بی‌سابقه‌ای در تعیین مسئولیت کیفری پدید آورده است، چراکه زنجیرهٔ تولید و انتشار محتوای جعلی شامل بازیگران متعددی از جمله توسعه‌دهندگان الگوریتم‌های مولد و گردآورندگان مجموعه‌داده‌های آموزشی، طراحان رابط‌های کاربری، ارائه‌دهندگان سرویس‌های ابری، کاربران نهایی و حتی خود سیستم‌های هوش مصنوعی خودمختار می‌شود. از دیدگاه جرم‌شناسی سایبری، این وضعیت به معضل شناسایی چندعاملی[1] منجر می‌شود (Korshunov and Marcel, 2019: 4) که در آن تعیین سهم هریک از عوامل در ارتکاب جرم به دلیل پیچیدگی فنی و ذاتی سیستم‌های یادگیری عمیق و نبود شفافیت در فرایندهای تصمیم‌گیری الگوریتمی، بسیار دشوار است (Gong and Li, 2024: 580).

برای مثال، در مواردی که الگوریتمی خودمختار با استفاده از داده‌های آموزشی غیرقانونی (مانند تصاویر بدون رضایت افراد)، محتوای جعلی تولید می‌کند، تعیین اینکه آیا مسئولیت اصلی بر عهدهٔ توسعه‌دهندهٔ الگوریتم است، یا مالک داده‌های آموزشی، یا کاربری که از سیستم سوءاستفاده کرده است، به تحلیل‌های پیچیدهٔ حقوقی و فنی نیاز دارد که هنوز چهارچوب نظری جامعی برای آن توسعه نیافته است. راهکارهای پیشنهادی در ادبیات آکادمیک شامل گسترش «چهارچوب‌های توزیع مسئولیت پویا»[2] است که بتواند با درنظرگرفتن معیارهایی مانند میزان کنترل هر بازیگر بر فرایند تولید محتوا، دانش فنی آن‌ها از قابلیت‌های سیستم، و میزان سوءنیت در استفاده از فناوری، سهم هر عامل را در ارتکاب جرم به‌صورت علمی محاسبه کند (حاجی اسماعیلی، ۱۴۰۳: ۹۵). همچنین، برقراری «نظام‌های ثبت و رهگیری الگوریتمی»[3] مبتنی بر فناوری‌های دفترکل توزیع‌شده می‌تواند شفافیت لازم را در زنجیرهٔ تولید و انتشار محتوای دیپ‌فیک فراهم آورد (Schuett, 2019: 13).

1. Multi-Agent Identification Problem
2. Dynamic Liability Distribution Frameworks
3. Algorithmic Provenance Systems

۱-۲. مشکلات در اثبات نیت و قصد جنایی

مسئلهٔ اثبات عنصر ذهنی[1] در جرایم مبتنی بر دیپ‌فیک یکی از پیچیده‌ترین چالش‌های حقوق کیفری در عصر دیجیتال محسوب می‌شود، چراکه ماهیت واسطه‌ای و غیرخطی سیستم‌های هوش مصنوعی مولد، ارتباط سنتی بین قصد مجرمانه و نتیجهٔ مجرمانه را بسیار متحول کرده است (جزایری و عظیمی، ۱۳۹۹: ۶). از منظر تئوری حقوقی، سیستم‌های دیپ‌فیک با «گسست شناختی»[2] در زنجیرهٔ علیت جنایی، امکان تطبیق مستقیم معیارهای سنتی اثبات نیت مجرمانه (مانند عمد، علم و بی‌مبالاتی) را با واقعیت‌های فنی این جرایم با مشکل مواجه می‌سازند. بر اساس مطالعات،[3] در ۷۲درصد از پرونده‌های مرتبط، دادگاه‌ها در تشخیص دانش فنی لازم برای احراز سوءنیت کاربران نهایی چالش داشته‌اند، به‌ویژه زمانی که کاربر از رابط‌های کاربری ساده‌ای استفاده می‌کند که پیچیدگی الگوریتم‌های زیربنایی را پنهان می‌سازد (Binns and Veale, 2021: 5–6).

این وضعیت به معضل مسئولیت شناختی منجر شده است، که در آن حتی در صورت وجود قصد اولیه در کاربر، ماهیت غیرشفاف فرایندهای یادگیری عمیق (مانند تبدیل‌های غیرخطی در لایه‌های پنهان شبکه‌های عصبی) امکان پیش‌بینی دقیق نتایج مجرمانه را ناممکن می‌سازد (ناصری و آهنگران، ۱۴۰۵: ۱۸۲). از دیدگاه جرم‌شناسی سایبری، این پدیده به بازتعریف معیارهای سنتی «احتمال معقول وقوع نتیجه»[4] در پرتو ویژگی‌های فنی سیستم‌های مولد نیاز دارد، زیرا بر اساس تحقیقات تجربی، رفتارهای ظاهراً سادهٔ کاربران (مانند آپلود یک تصویر) ممکن است به نتایج پیچیده‌ای منتهی شود که حتی توسعه‌دهندگان متخصص نیز قادر به پیش‌بینی کامل آن‌ها نباشند (نوری میکائیل آباد، ۱۴۰۳: ۶۹).

۱-۳. مستندات جعلی و نقش آن در بروز نارسایی سیستم قضایی

از پیچیده‌ترین چالش‌های نظام‌های قضایی معاصر، مواجهه با پدیدهٔ مستندات

---

1. Mens Rea

۲. گسست شناختی در دیپ‌فیک به وضعیتی اشاره دارد که در آن نظام ادراکی و قضاوت ذهنی انسان در تشخیص مرز میان واقعیت و جعل دچار اختلال می‌شود، زیرا محتوای تولیدشده از طریق دیپ‌فیک با الگوهای شناختی پیشین فرد دربارهٔ «اصالت تصویر، صدا یا ویدئو» سازگار به نظر می‌رسد، اما در واقع فاقد منشأ واقعی است.

۳. آمار رسمی و تفکیکی منتشرشده از دیپ‌فیک در ایران (مثلاً از سوی نهادهای قضایی یا پلیس) به‌صورت عمومی منتشر نشده است. تحقیقات حقوقی داخلی هم در خصوص دیپ‌فیک وجود دارد، اما رأی قضایی رسمی که در پروندهٔ کیفری واقعی دیپ‌فیک صادر شده باشد، یافت نشده است. از طرفی می‌دانیم که دیپ‌فیک به‌عنوان مصداق خاص در قانون کیفری ایران جرم‌انگاری مستقیم نشده و بیشتر تحت عناوین کلی مانند جعل رایانه‌ای، نشر اکاذیب یا هتک حیثیت قابل پیگرد است؛ از این رو، بررسی آرای قضایی و ارائهٔ آمار در خصوص چالش‌برانگیزبودن استفاده از دیپ‌فیک در جرایم در مرحلهٔ رسیدگی قضایی، ممکن به نظر نمی‌رسد.

4. Foreseeability Doctrine

جعلی[۱] تولیدشده از طریق هوش مصنوعی[۲] است که به دلیل پیشرفت‌های چشمگیر در الگوریتم‌های مولد مانند شبکه‌های عصبی مولد تخاصمی و مدل‌های انتشار، قادر به تولید محتوای بصری و صوتی با سطحی از واقع‌نمایی‌اند که حتی کارشناسان خبره نیز به سادگی نمی‌توانند تشخیص دهند. بر اساس مطالعات تجربی در حوزهٔ پزشکی قانونی دیجیتال، نرخ خطای تشخیصی در تمایز بین محتوای واقعی و دیپ‌فیک در آزمایش‌های کنترل‌شده به ۴۷درصد می‌رسد، که این مسئله تهدیدی جدی برای اصل قابلیت اعتماد در نظام ادلهٔ الکترونیکی محسوب می‌شود. از منظر حقوقی، این وضعیت به «معضل اعتبارسنجی چندلایه» منجر شده است، چراکه شواهد دیپ‌فیک علاوه بر مرحلهٔ جمع‌آوری ادله، در مراحل تحلیل، نگهداری و ارائه در دادگاه نیز می‌توانند اعتبار نظام دادرسی را تحت تأثیر قرار دهند (سی.کینگستون و همکاران، ۱۴۰۲: ۱۱۶–۱۱۸). طبق تحقیقات، در ۶۸درصد پرونده‌های مرتبط، قضات در ارزیابی وزن ادلهٔ دیجیتالی با چالش‌های بنیادینی مواجه بوده‌اند، به‌ویژه زمانی که این ادله با استفاده از الگوریتم‌های یادگیری عمیق تولید شده‌اند که فاقد «زنجیرهٔ سرپرستی دیجیتال» استاندارد هستند (Chac, 2020: 25).[۳]

۲. راهکارها و پیشنهادها

با توجه به دشواری‌های گوناگون ناشی از جرایم مبتنی بر دیپ‌فیک، ارائهٔ راهکارهای جامع و کاربردی برای بهبود چهارچوب‌های قانونی و قضایی موضوعی مهم به شمار می‌رود. در ادامه، چندین روش پیشنهادی برای رفع مشکلات موجود و همگام‌سازی قوانین با فناوری‌های نوین بیان می‌شود.

۲-۱. تدوین و توسعهٔ قوانین اختصاصی دیپ‌فیک

طراحی قوانین ویژه برای جرایم مبتنی بر دیپ‌فیک ضروری است، زیرا این فناوری با سرعتی بالا در حال تحول است و خلأهای قانونی چشمگیری پدید آورده است (رحیمی قدیم و طبسی، ۱۴۰۳: ۷). نخستین گام، تعریف دقیق جرم دیپ‌فیک است که شامل هرگونه دستکاری غیرمجاز در محتوای صوتی، تصویری یا ویدئویی با هدف فریب مخاطب یا آسیب به اشخاص حقیقی و حقوقی باشد.

---

۱. مستندات جعلی به اصطلاح انگلیسی «Fabricated Evidence» منظور مستندات دروغین ساخته‌شده از طریق هوش مصنوعی است.

2. AI-Generated Evidence

۳. از آنجا که آمار رسمی در خصوص چالش‌های قضایی دیپ‌فیک در ایران یافت نشده است، لذا در این پژوهش برای بیان چالش‌های حقوقی در مرحلهٔ رسیدگی قضایی که ممکن است سیستم قضایی کشور را با مشکل مواجه کند، از آمارهای مدون نمونه‌های خارجی استفاده شده است.

این تعریف باید مرز میان استفاده مشروع و غیرمشروع از دیپ‌فیک را روشن کند و مسئولیت تولید، انتشار و استفادهٔ غیرقانونی از محتواهای جعل‌شده را مشخص کند. مسئولیت‌پذیری می‌تواند سلسله‌مراتبی باشد؛ تولیدکنندگان نرم‌افزار دیپ‌فیک، توزیع‌کنندگان محتوای جعلی و کاربرانی که برای اهداف مجرمانه از آن استفاده می‌کنند، هریک مطابق نقش خود پاسخ‌گو باشند. چنین رویکردی هم از تکرار جرم جلوگیری می‌کند و هم توسعه‌دهندگان را به گنجاندن سازوکارهای تشخیص و پیشگیری در محصولات خود ترغیب می‌کند.

۲-۲. پذیرش نظریهٔ مسئولیت محض برای تولیدکنندگان و توسعه‌دهندگان

با توجه به رفتارهای پیش‌بینی‌نشدهٔ دیپ‌فیک و امکان سوءاستفاده در حوزه‌های مالی و اجتماعی، اعمال نظریهٔ «مسئولیت محض» برای تولیدکنندگان و توسعه‌دهندگان این فناوری می‌تواند بازدارنده باشد (پارسا، ۱۴۰۲: ۳۵۷). بر اساس این نظریه، بدون نیاز به اثبات تقصیر یا قصد مجرمانه، تولیدکننده یا توسعه‌دهنده مسئول پیامدهای محتوای جعلی است. این رویکرد انگیزه‌ای قوی برای رعایت استانداردهای امنیتی و تضمین صحت محتوای تولیدی ایجاد می‌کند و در عین حال می‌تواند با سازوکارهای بیمه‌ای همراه شود تا فشار مالی بر توسعه‌دهندگان کاهش یابد.

۲-۳. الزام به شفافیت و افشای عملکرد الگوریتم‌های دیپ‌فیک

وضع و اجرای مقررات الزام‌آور برای افشای جزئیات عملکرد الگوریتم‌های دیپ‌فیک و داده‌های آموزشی مورد استفاده در تولید محتوای جعلی، راهکاری کلیدی در خصوص افزایش شفافیت و تقویت پاسخ‌گویی در حوزهٔ حقوقی محسوب می‌شود. این الزامات به نهادهای قضایی، کارشناسان فنی و مراجع تحقیقاتی امکان می‌دهد تا در مواجهه با پرونده‌های مبتنی بر جعل عمیق، منشأ محتوا، نحوهٔ تولید و افراد دخیل در فرایند ساخت محتوای جعلی را به‌صورت دقیق شناسایی و پیگیری کنند. اهمیت این رویکرد به‌ویژه در پرونده‌هایی که اهداف اقتصادی، سیاسی یا اجتماعی دارند، برجسته می‌شود، زیرا نبود شفافیت در تولید محتوای دیپ‌فیک می‌تواند به پیچیدگی‌های جدی در تعیین مسئولیت کیفری و حقوقی منتهی شود و امکان احراز سوءنیت یا تقصیر مجرمان را کاهش دهد.

یکی از ابزارهای فنی مؤثر در این زمینه، واترمارکینگ دیجیتال است. این روش با تعبیهٔ نشانه‌های دیجیتال حذف‌نشدنی در محتوای صوتی، تصویری یا ویدئویی، امکان شناسایی منبع اولیهٔ محتوا و ردیابی مسیر انتشار آن را فراهم می‌کند. این راهکار مبتنی بر اصول رمزنگاری و پردازش سیگنال عمل می‌کند و می‌تواند به‌منزلهٔ استانداردی عملی در تولید محتوای دیپ‌فیک به کار گرفته شود تا از انتشار گستردهٔ محتوای جعلی و سوءاستفاده‌های احتمالی جلوگیری شود.

علاوه بر این، اســتفاده از واترمارکینگ دیجیتال به کارشناســان قضایی این امکان را می‌دهد که در مرحلهٔ تحقیق و تفحص، ارتباط بین تولیدکننده، توزیع‌کننده و کاربر نهایی محتوا را با دقت بیشتری مستندسازی کنند و بدین ترتیب از وقوع اشتباهات قضایی و تضییع حقوق قربانیان جلوگیری شود (میرکمالی و همکاران، ۱۴۰۵: ۱۱۷).

از منظر حقوقی و سیاست‌گذاری، الزام به افشای جزئیات فنی الگوریتم‌ها و داده‌های آموزشی باید با رعایت حریم خصوصی، اســرار تجاری و امنیت ملی تعادل یابد تا ضمن ارتقای شــفافیت، هیچ‌یک از حقوق مشروع شرکت‌ها یا کاربران نادیده گرفته نشود. سازوکارهای نظارتی باید قابلیت راســتی‌آزمایی اطلاعات افشاشــده را داشــته باشــند و در صورت تخلف، جریمه‌های متناسب و ســازوکارهای اصلاحی اعمال شود. تجربیات موفق در حوزه‌هایی مانند مقررات عمومی حفاظت از داده‌ها[1] نشــان می‌دهد که چنین الزاماتی می‌توانند ضمن افزایش اعتماد عمومی، اســتانداردهای فنی و عملیاتی تولید محتوا را بهبود ببخشند و زمینه را برای گسترش مسئولانه و پایدار فناوری‌های دیپ‌فیــک فراهــم کنند. در نتیجه، وضــع مقررات الزام‌آور برای افشــای عملکــرد الگوریتم‌های دیپ‌فیک، علاوه بر آنکه ابزار مهمی برای پیشگیری و کاهش سوءاستفاده‌های قانونی و غیرقانونی اســت، می‌تواند به ایجاد زیرساخت‌های قضایی و کارشناسی قوی‌تر در مواجهه با جرایم مبتنی بر جعــل عمیق کمک کنــد و در نهایت، موجب ارتقای اعتماد عمومی و مشــروعیت نظام قضایی در مواجهه با محتوای دیجیتال شود (Garvey, 2020: 7).

۲-۴. گسترش آموزش و تخصص قضایی در دیپ‌فیک

قضات، کارشناســان و ضابطان قضایی باید آموزش‌های ویژه‌ای در شناســایی و ارزیابی محتوای دیپ‌فیک دریافت کنند. این آموزش‌ها شــامل شــناخت ویژگی‌های فنی دیپ‌فیک، ارزیابی اعتبار شــواهد دیجیتال و کاربرد ابزارهای تشخیص معتبر اســت. توانمندسازی نهادهای قضایی در این حوزه، تضمین‌کنندهٔ دادرسی منصفانه و کاهش خطاهای قضایی در پرونده‌های مبتنی بر جعل عمیق خواهد بود (O'Keefe, 2025: 2).

نتیجه‌گیری

در ایــن پژوهش چالش‌های حقوقی ناشــی از جرایم مبتنی بر جعل عمیــق (دیپ‌فیک) در مرحلهٔ رســیدگی قضایی بررســی شد. بر اساس یافته‌ها، این فناوری‌های پیشــرفته، نظام‌های قضایی را با چالش‌های بی‌ســابقه‌ای در تعیین مســئولیت کیفری، اثبات عنصر ذهنی و اعتبارسنجی ادله مواجه

1. GDPR

کرده‌اند. در تحلیل‌های انجام‌شده تأکید شده است که بهره‌گیری صرف از رویکردهای سنتی حقوق کیفری برای پاسخ‌گویی به این چالش‌های پیچیده کافی نیست، بلکه گسترش چهارچوب‌های نوین حقوقی و فنی ضرورت دارد. این چهارچوب‌ها باید بتوانند با ترکیب روش‌های پیشرفتهٔ اعتبارسنجی، سیستم‌های نظارتی هوشمند و استانداردهای بین‌المللی، تعادلی میان پیشرفت فناوری و امنیت حقوقی برقرار کنند. در این راستا، چندین راهکار مهم شناسایی شدند که عبارت‌اند از تدوین قوانین اختصاصی برای جرایم مبتنی بر دیپ‌فیک، اعمال نظریهٔ مسئولیت محض برای تولیدکنندگان، الزام به شفاف‌سازی عملکرد سیستم‌ها از طریق روش‌هایی مانند واترمارکینگ، گسترش آموزش و تخصص قضایی در دیپ‌فیک. این راهکارها علاوه بر آنکه باعث کاهش جرایم و مخاطرات مرتبط با دیپ‌فیک می‌شوند، زمینه را برای گسترش مسئولانه و پایدار این فناوری فراهم می‌سازند.

در سطح کلان، موفقیت این راهکارها وابسته به همکاری گسترده و چندجانبه بین نهادهای قانون‌گذاری، جامعهٔ علمی و صنعت فناوری در عرصهٔ بین‌المللی است. این همکاری باید به‌گونه‌ای طراحی شود که از یک سو با ایجاد استانداردهای بین‌المللی هماهنگ، از شکاف‌های نظارتی جلوگیری کند و از سوی دیگر، با حفظ انعطاف‌پذیری لازم، مانع سرکوب نوآوری‌های مفید نشود. بر اساس پژوهش حاضر، ترکیب هوشمندانهٔ اصول حقوقی، فناوری‌های پیشرفتهٔ نظارتی و چهارچوب‌های اخلاقی می‌تواند به برقراری اکوسیستم دیجیتالی امن‌تر و عادلانه‌تری منجر شود، که در آن هم از حقوق افراد محافظت می‌شود و هم فضای لازم برای رشد فناوری‌های تحول‌آفرین فراهم می‌شود. این رویکرد یکپارچه و نظام‌مند، می‌تواند به‌منزلهٔ الگویی برای مواجهه با چالش‌های پیچیدهٔ فناوری‌های نوین در آینده استفاده شود و مسیر را برای تحقیقات آتی در این حوزهٔ مهم هموار سازد. در نتیجه، در این پژوهش بر ضرورت گسترش زیرساخت‌های حقوقی و فنی هماهنگ با سرعت پیشرفت فناوری تأکید شده و نشان داده شده است که فقط از طریق تعامل سازنده بین تمام ذی‌نفعان می‌توان به راهکارهای مؤثر و پایدار در این زمینه دست یافت.

منابع

۱. ابراهیمی، علیرضا. (۱۴۰۲). **کاربرد هوش مصنوعی در امور حقوقی: فرصت‌ها و چالش‌ها**، مطالعات حقوق، دورهٔ ۳۵، ۲۳-۳۴.

۲. پارسا، ناهید. (۱۴۰۲). **مسئولیت مدنی در خودروهای تمام‌خودران در حقوق ایران و آلمان با نگاهی به قانون خودروهای خودران ۲۰۲۱ آلمان**، حقوق خصوصی، شمارهٔ ۴۳، ۳۴۸-۳۶۸.

۳. جزایری، سیدعباس و عظیمی، مینا. (۱۳۹۹). **بررسی مسئولیت کیفری هوش مصنوعی در حقوق کیفری ایران**، ششمین کنفرانس ملی علوم انسانی و مطالعات مدیریت، تهران.

۴. حاجی اسماعیلی، میلاد. (۱۴۰۳). **چالش‌های مسئولیت مدنی هوش مصنوعی در نظام حقوقی ایران؛ با نگاهی به مقررات‌گذاری در اتحادیهٔ اروپا**، فصلنامه دولت و حقوق، دورهٔ ۵، شمارهٔ ۱، ۸۱–۹۸.
۵. رحیمی قدیم، علیرضا و طبسی، بهاره. (۱۴۰۳). **تطبیق قوانین حقوقی با پیشرفت‌های فناوری در صنعت مهندسی**، اولین همایش ملی تأثیر متقابل حقوق بین‌الملل و حقوق داخلی در توسعهٔ قوانین.
۶. رضایی، سارا و امینی، محمد. (۱۴۰۳). **مسئولیت کیفری در دیپ‌فیک و جعل هویت تصویری**، مجله حقوق سایبری، دورهٔ ۱، شمارهٔ ۲، ۱–۲۰.
۷. سی. کینگستون، جی. کی؛ صلح‌چی، سارا و بیگلربیگی، کیان. (۱۴۰۲). **هوش مصنوعی و مسئولیت قانونی**، تمدن حقوقی، دورهٔ ۶، شمارهٔ ۱۸، ۱۰۳–۱۲۰.
۸. علی‌ین، آرش. (۱۴۰۰). **جرم‌انگاری دیپ‌فیک‌ها از منظر تعهدات حقوق بشری دولت‌ها**، دوفصلنامه تحقیقات حقوق قضایی، دورهٔ ۲، شمارهٔ ۳، ۱–۱۸.
۹. فاخری، سهیل؛ نوربخش، اعظم‌السادات و یمقانی، محمدرضا. (۱۴۰۳). **مدل‌ها و روش‌های تشخیص دیپ‌فیک در هوش مصنوعی و تأثیر این پدیده بر فرهنگ اجتماعی از منظر رسانه**، فصلنامه مدیریت نوآوری و راهبردهای عملیاتی، دورهٔ ۵، شمارهٔ ۳، ۱–۲۲.
۱۰. فرزین، امید و سمیعی، روح‌الله. (۱۴۰۲). **چالش‌های اخلاقی و حقوقی استفاده از هوش مصنوعی در کسب‌وکارهای دیجیتال**، تکنولوژی در کارآفرینی و مدیریت استراتژیک، دورهٔ ۲، شمارهٔ ۲، ۱۷–۲۷.
۱۱. فلاح تفتی، فاطمه و عبدخدایی، زهره. (۱۴۰۴). **تنظیم‌گری دیپ‌فیک در پرتو فقه حکومتی: الگوی فقهی ــ حقوقی برای حکمرانی بر این فناوری**، دوفصلنامه حقوق فناوری‌های نوین، دورهٔ ۶، شمارهٔ ۱۲، ۱–۲۵.
۱۲. میرکمالی، سیدعلیرضا، حیدرپور، حمیدرضا و مهرآرا، ژیلا. (۱۴۰۵). **تحلیل ضرورت جرم انگاری افتراقی جعل عمیق در نظام کیفری ایران (در پرتو مطالعه رویکرد حقوق کیفری آمریکا و چین)**، فصلنامه قضاوت، ۲۶(۱۲۵)، ۱۰۴–۱۲۵. 1540. doi: 10.22034/judg.2026.2066883 .
۱۳. ناصری، جمال الدین و آهنگران، محمدرسول . (۱۴۰۵). **هوش مصنوعی به مثابه کارشناس رسمی دادگستری**، فصلنامه قضاوت، ۲۶(۱۲۶)، ۱۷۱–۱۹۱. doi: 10.22034/judg.2025.2055221.1390 .
۱۴. نوری میکائیل آباد، امیرمهدی. (۱۴۰۳). **نقش هوش مصنوعی در فرایندهای تصمیم‌گیری**، مطالعات کاربردی در علوم مدیریت و توسعه، شمارهٔ ۵۴، ۶۵–۷۲.

15. Ahmed, S. R.; Sonuç, E.; Ahmed, M. R. and Duru, A. D. (2022). **"Analysis survey on deepfake detection and recognition with convolutional neural networks"**, Int. Congr. Hum. -Comput. Interact., Optim. Robot. Appl. (HORA), pp. 1-7.
16. Binns, R. and Veale, M. (2021). **"Fairness in Machine Learning: Lessons from Political Philosophy"**, In Proceedings of the 38th International Conference on Machine Learning, PMLR.
17. Chac, Y. (2020). **"U.S. AI Regulation Guide: Legislative Overview and Practical Considerations"**, The Journal of Robotics, Artificial Intelligence & Law, Vol. 3, No. 1, pp. 17-40.
18. Garvey, J. B. (2022). **"Let's Get Real: Weak Artificial Intelligence Has Free Speech Rights"**, Fordham L. Rev., No. 91.
19. Gong, L. Y. and Li, X. J. (2024). **"A contemporary survey on deepfake detection: datasets, algorithms, and challenges"**, Electronics, No. 13.
20. Helm, J. M., et al. (2020). **"Machine Learning and Artificial Intelligence: Definitions, Applications, and Future Directions"**, The Use of Technology in Orthopaedic Surgery Intraoperative and Post-Operative Management (C Krueger and S Bini, Section Editors), Vol. 13, 69-76.
21. Korshunov, P. and Marcel, S. (2019). **"Vulnerability assessment and detection of deepfake videos"**, In 2019 International Conference on Biometrics (ICB), IEEE (2019, June), pp. 1-6.

22. Marr, B. and Ward, M. (2019). **Artificial Intelligence in Practice**, United Kingdom: John Wiley and Sons Ltd.
23. Nayak, J.; Naik, B.; Vimal, S. and Favorskaya, M. (2024). **Machine Learning for Cyber Physical System: Advances and Challenges**, Germany: Springer.
24. O'Keefe, C.; Ramakrishnan, K.; Tay, J. and Winter, C. (2025). **"Law-Following AI: designing AI agents to obey human laws"**, Forthcoming in the 94th volume of the Fordham Law Review.
25. Schuett, J. (2019). **"A Legal Definition of AI"**, arXiv preprint arXiv:1909.01095, 1-14.
26. Wachter, S. (2024). **"Limitations and Loopholes in the EU AI Act and AI Liability Directives: What This Means for the European Union, the United States, and Beyond"**, Yale Journal of Law & Technology, Vol. 26, No. 3, pp. 671-718.
27. Walters, R. and Novak, M. (2021). **Cyber Security, Artificial Intelligence, Data Protection & the Law**, Germany: Springer.