Modern Interdisciplinary Research in Law
Vol. 6, Issue 2, Summer 2026, 105-119
jmirl.apadana.ac.ir

Apadana Institute of Higher Education

Criminal Liability Arising from Errors in Facial Recognition Systems in Criminal Cases

Hedieh Afroozi — M.A. Student in Criminal Law and Criminology, Lahijan Branch, Islamic Azad University, Lahijan, Iran

Amirreza Mahmoudi* — Assistant Professor, Department of Law, Lahijan Branch, Islamic Azad University, Lahijan, Iran

Received: 01/June/2026 Accepted: 15/June/2026 eISSN: 2783-4204 ISSN: 2783-3631

Abstract

Although the deployment of facial recognition systems in criminal cases has accelerated crime detection, linking justice to intelligent algorithms has created unprecedented challenges in the realm of fundamental rights and the foundations of liability. Using a descriptive–analytical method, the present study seeks to examine the legal nature of errors produced by these systems and addresses the question of which legal model can ground the attribution of criminal liability when misidentification occurs and individual liberties are violated. The findings indicate that, due to the technical complexity commonly described as the "black box" in deep-learning layers and the presence of structural "biometric biases," the outputs of these systems lack certainty and operate only at the level of "mathematical probabilities." From a legal perspective, the systems' relative autonomy generates a "causal gap" in attributing the offense and places the classical, human-centered concepts of "fault" and "will" in a dead end. In Iran's criminal justice system, the absence of explicit provisions in the Computer Crimes Law and the lack of an independent legal personality for artificial intelligence have resulted either in unfair liability for users or in de facto immunity for negligent designers. The study concludes that, in order to safeguard the presumption of innocence, the outputs of this technology should not be accepted as "independent evidence"; rather, they should be treated only as a form of "judicial presumption" and remain subject to direct human oversight. Moreover, based on a "distributed liability" approach, it is necessary to identify the chain of responsibility between the designer and the operator, and where fault is ambiguous, to rely on the state's compensatory liability—pursuant to Article 171 of the Constitution—to compensate victims for resulting harm.

Keywords: facial recognition technology (FRT), criminal liability, artificial intelligence, algorithmic error, biometric bias, fundamental rights

* Corresponding Author: Dr.mahmoudi@iau.ac.ir

فصلنامه تحقیقات نوین میان‌رشته‌ای حقوق
دوره ٦، شماره ٢، تابستان ١٤٠٥، ١٠٥-١١٩
jmirl.apadana.ac.ir

تاریخ دریافت: ۱۴۰۵/۰۳/۱۱ تاریخ پذیرش: ۱۴۰۵/۰۳/۲۵ eISSN: 2783-4204 ISSN: 2783-3631

مسئولیت کیفری ناشی از اشتباهات سیستم‌های تشخیص چهره در پرونده‌های جنایی

هدیه افروزی دانشجوی کارشناسی ارشد حقوق کیفری و جرم‌شناسی، دانشگاه آزاد اسلامی، لاهیجان، ایران

امیررضا محمودی* استادیار گروه حقوق، دانشگاه آزاد اسلامی، لاهیجان، ایران

چکیده

استقرار سیستم‌های تشخیص چهره در پرونده‌های جنایی اگرچه سرعت کشف جرم را ارتقا بخشیده، اما پیوند عدالت با الگوریتم‌های هوشمند، چالش‌های بی‌سابقه‌ای را در حوزه‌ی حقوق بنیادین و مبانی مسئولیت پدید آورده است. پژوهش حاضر با هدف واکاوی ماهیت حقوقی اشتباهات ناشی از این سامانه‌ها، با روشی توصیفی-تحلیلی به بررسی این پرسش می‌پردازد که در صورت بروز خطای شناسایی و نقض آزادی‌های فردی، بار مسئولیت کیفری بر مبنای کدام الگوی حقوقی قابل انتساب است. یافته‌های پژوهش نشان می‌دهند که به‌دلیل پیچیدگی‌های فنی موسوم به «جعبه سیاه» در لایه‌های یادگیری عمیق و وجود «سوگیری‌های بیومتریک» ساختاری، خروجی این سیستم‌ها فاقد وصف قطعیت بوده و صرفاً در تراز «احتمالات ریاضی» قرار دارد. از منظر حقوقی، استقلال نسبی این سامانه‌ها منجر به ایجاد «خلأ سببی» در انتساب بزه شده و مفاهیم کلاسیک «تقصیر» و «اراده» را که بر مدار انسان‌محوری استوارند، با بن‌بست مواجه ساخته است. در نظام کیفری ایران، فقدان مقررات صریح در قانون جرایم رایانه‌ای و نبود شخصیت حقوقی مستقل برای هوش مصنوعی، منجر به ایجاد مسئولیت‌های ناعادلانه برای کاربران یا مصونیت طراحان متقصر شده است. نتیجه‌گیری پژوهش مؤید آن است که برای صیانت از اصل برائت، خروجی‌های این فناوری نباید به‌عنوان «دلیل مستقل» پذیرفته شوند، بلکه باید صرفاً منزلت «اماره قضایی» یافته و تحت نظارت مستقیم انسانی قرار گیرند. همچنین، بر اساس رهیافت «مسئولیت توزیعی»، ضرورت دارد زنجیره‌ی مسئولیت میان طراح و اپراتور شناسایی شده و در موارد ابهام در تقصیر، از ظرفیت مسئولیت جبرانی دولت بر مبنای اصل ۱۷۱ قانون اساسی جهت جبران خسارت بزه‌دیدگان استفاده گردد.

کلیدواژه‌ها: تشخیص چهره (FRT)، مسئولیت کیفری، هوش مصنوعی، خطای الگوریتمی، سوگیری بیومتریک، حقوق بنیادین

* نویسندهٔ مسئول: Dr.mahmoudi@iau.ac.ir

مقدمه

عصر حاضر دوران گذار از دادرسی‌های سنتی به «عدالت دیجیتال» است؛ فرایندی که در آن فناوری‌های نوین نه‌تنها ابزارهای کمکی بلکه به کنشگران کلیدی در اتخاذ تصمیمات قضایی بدل شده‌اند. تحولات شگرف فناوری در دهه اخیر، پارادایم‌های سنتی کشف جرم و شناسایی مجرمان را با دگرگونی بنیادین مواجه ساخته است. در این میان، هوش مصنوعی با نفوذ به لایه‌های مختلف پلیس علمی، وعده‌ی استقرار امنیت هوشمند را می‌دهد. شاخص‌ترین جلوه‌ی این تحول، ورود هوش مصنوعی به حوزه‌ی تحلیل داده‌های بیومتریک، به‌ویژه «سیستم‌های تشخیص چهره»، در فرایندهای جنایی است. این فناوری با بهره‌گیری از الگوریتم‌های یادگیری ماشین، نه‌تنها حجم عظیمی از داده‌های تصویری را با سرعتی فراتر از توان انسانی تحلیل می‌کند، بلکه به‌عنوان بازوی کمکی و مکمل تشخیصی ضابطان قضایی شناخته می‌شود.

با این حال، این «شیفتگی تکنولوژیک» نباید منجر به نادیده انگاشتن لایه‌های تاریک و مبهم الگوریتم‌ها گردد. در کنار این کارکردهای غیرقابل انکار، نباید از پیامدهای مخاطره‌آمیز ناشی از خطاهای احتمالی این سامانه‌ها غافل ماند. چالش اصلی زمانی رخ می‌نماید که فناوری تشخیص چهره، به‌جای «تسهیل عدالت» به «تولید اشتباه قضایی» منجر شود. گزارش‌های متعدد از تشخیص‌های نادرست و بازداشت‌های ناعادلانه مبتنی بر خروجی‌های خودکار، زنگ خطری جدی برای نظام‌های حقوقی است تا در مبانی سنتی مسئولیت کیفری بازنگری کنند (کفانی‌نژاد، ۱۴۰۴، ص۲).

بحران پیش‌رو تنها یک نقیصه‌ی فنی نیست بلکه یک «بن‌بست حقوقی» است؛ چراکه مفاهیم کلاسیک «تقصیر» و «اراده» با ظهور موجودیت‌های نیمه‌خودمختار، کارآیی خود را از دست داده‌اند. از این‌رو، پرسش بنیادین اینجا است که در صورت بروز اشتباه در شناسایی و متعاقب آن نقض حقوق بنیادین اشخاص، بار مسئولیت کیفری چگونه توزیع می‌شود؟ آیا مقام انسانی همچنان تنها پاسخ‌گو است یا باید بر اساس نظریات نوین مسئولیت، قائل به توزیع مسئولیت میان توسعه‌دهندگان نرم‌افزار، سیاست‌گذاران و اپراتورهای نهایی بود؟ (کفانی‌نژاد، ۱۴۰۴، ص ۳). پژوهش حاضر در صدد است تا با واکاوی مبانی فنی و حقوقی این سامانه‌ها، مدل مطلوبی از انتساب مسئولیت را در نظام کیفری ایران ترسیم نماید.

به نظر می‌رسد الگوی سنتی مسئولیت کیفری که بر مدار انسان‌محوری استوار است، در مواجهه با پدیده «جعبه سیاه» و استقلال نسبی الگوریتم‌ها، کارایی کافی نداشته و منجر به ایجاد «خلأ سببی» در انتساب بزه می‌گردد. از این‌رو، فرضیه‌ی اصلی پژوهش بر این است که با پذیرش نظریه‌ی «مسئولیت توزیعی» میان زنجیره کنشگران (طراح و اپراتور) و تنزل جایگاه اثباتی این سامانه‌ها از «دلیل مستقل» به «اماره قضایی»، می‌توان ضمن صیانت از اصل برائت، از ظرفیت مسئولیت جبرانی دولت بر مبنای اصل ۱۷۱ قانون اساسی برای جبران خسارات ناشی از خطاهای سیستمی بهره جست.

پیشینه پژوهش

پژوهش در باب تلاقی «فناوری‌های نوین» و «حقوق کیفری» در سالیان اخیر شتاب فزاینده‌ای گرفته است. با این حال، واکاوی تخصصی مسئولیت ناشی از خطاهای الگوریتمی در سامانه‌های تشخیص چهره، حوزه‌ای نوپدید محسوب می‌شود. پیشینه‌ی مطالعاتی در این زمینه را می‌توان در سه محور داخلی و بین‌المللی دسته‌بندی نمود:

۱. مطالعات معطوف به ماهیت فنی و خطاهای بیومتریک: در بُعد فنی، سیگاری و همکاران (۱۳۹۱) در پژوهشی تحت عنوان «بازشناسی چهره در شرایط محیطی متغیر»، به تحلیل ناپایداری‌های الگوریتمی در مواجهه با تغییرات نوری و فیزیکی پرداخته‌اند. یافته‌های آنان مؤید این است که نرخ خطای سیستم‌های پردازش تصویر در شرایط واقعی به

مراتب بالاتر از محیط‌های آزمایشگاهی است. در سطح بین‌المللی نیز جکسون[1] (۲۰۱۹) در گزارش‌های انتقادی خود، بر عدم قطعیت خروجی‌های این سامانه‌ها تأکید ورزیده و آن‌ها را نه یک «دلیل اثباتی»، بلکه صرفاً یک «سرنخ تحقیقاتی» متزلزل معرفی کرده است که پتانسیل بالایی برای تولید اشتباهات قضایی دارند.

۲. مطالعات مرتبط با مبانی مسئولیت کیفری هوش مصنوعی: در حوزه‌ی مسئولیت، الجیزانی و حیدری (۱۴۰۴) در نوشتاری به بررسی چالش «اراده و فاعلیت» در سیستم‌های مستقل پرداخته و معتقدند که ساختار سنتی حقوق ایران که بر مدار «انسان‌محوری» استوار است، توان درک استقلال الگوریتمی را ندارد. همچنین، زندی و رفیعی علوی (۱۴۰۳) با نگاهی فلسفی-حقوقی، تفاوت میان «اراده الگوریتمی» و «اراده انسانی» را تحلیل کرده و ضرورت گذار از تفاسیر کلاسیک به سمت وضع قواعد نوظهور را تبیین نموده‌اند. در رویکردی تطبیقی، آلمیدا و همکاران[2] (۲۰۲۲) به بررسی سوگیری‌های نژادی و جنسیتی در الگوریتم‌ها پرداخته و بر لزوم «پاسخگویی الگوریتمی» در نظام‌های دادرسی تأکید کرده‌اند.

۳. مطالعات معطوف به سیاست جنایی و حقوق دادرسی: در قلمرو سیاست جنایی ایران، نظری و همکاران (۱۴۰۰) بر لزوم نظارت مشارکتی و حکمرانی فضای سایبر تأکید داشته‌اند تا از اتکای انحصاری نهادهای رسمی به هوش مصنوعی کاسته شود. لطفی (۱۳۹۸) نیز در بحث مسئولیت رسانه‌های الکترونیک، مدل «مسئولیت نیابتی و جمعی» را پیشنهاد داده است که می‌تواند الگویی برای انتساب مسئولیت به مدیران پلتفرم‌های هوشمند باشد. همچنین، رحماندوست (۱۳۹۹) با تمرکز بر اصل ۱۷۱ قانون اساسی، مسئولیت جبرانی دولت را در قبال اشتباهات سیستماتیک قضایی مورد واکاوی قرار داده است.

نوآوری پژوهش حاضر: علی‌رغم تلاش‌های صورت گرفته، اغلب پژوهش‌های پیشین یا صرفاً بر «ابعاد فنی» تمرکز داشته‌اند و یا به‌صورت کلی به «هوش مصنوعی» پرداخته‌اند. نوآوری و تمایز پژوهش حاضر در این است که به‌طور اختصاصی بر «زنجیره انتساب مسئولیت در اشتباهات خاص تشخیص چهره» متمرکز شده و با تلفیق چالش «جعبه سیاه» و «سوگیری‌های بیومتریک»، به ارائه‌ی راهکارهای عملیاتی بر اساس مبانی حقوقی ایران (از جمله نظریه تقصیر فنی و مسئولیت جبرانی دولت) می‌پردازد؛ موضوعی که در مطالعات پیشین به‌صورت یکپارچه مورد مداقه قرار نگرفته است.

۱- مبانی فنی و گونه‌شناسی اشتباهات[3]

هوش مصنوعی در رهیافتی کلان، به توسعه و تکوین سامانه‌هایی معطوف است که از قابلیت بازتولید فرایندهای شناختی انسانی، به‌ویژه «یادگیری پویا» و «اتخاذ تصمیمات مبتنی بر داده»، برخوردارند. در بستر فرایندهای جنایی، الگوریتم‌های «بینایی ماشین»[4] از طریق آموزش بر روی پایگاه‌داده‌های عظیم تصویری، توانمندی شناسایی و استخراج الگوهای پیچیده چهره را کسب می‌کنند. با این حال، کارکرد این سیستم‌ها به‌طور مستقیم و حیاتی وابسته به کیفیت و جامعیت «داده‌های آموزشی»[5] است؛ به‌گونه‌ای که هرگونه نقص، عدم توازن یا ماهیت مغرضانه در این داده‌ها، می‌تواند منجر به بروز خطاهای فاحش و سوگیری‌های الگوریتمی گردد (کفانی‌نژاد، ۱۴۰۴، ص ۴). علاوه بر این، محدودیت‌های ذاتی این فناوری در درک زمینه‌های محیطی[6] و حساسیت مفرط به کیفیت تصاویر ورودی، ریسک تولید خروجی‌های نادرست و در نتیجه، مخاطره‌ی انتساب اشتباه جرم به افراد بی‌گناه را به شدت افزایش می‌دهد (کفانی‌نژاد، ۱۴۰۴، ص ۵).

1. Jackson
2. Almeida et al.
3. technical foundations & error typology
4. machine vision
5. training data
6. contextual understanding

دقت عملیاتی سامانه‌های پردازش تصویر چهره به‌شدت تحت تأثیر متغیرهای محیطی و فیزیکی قرار دارد؛ به‌گونه‌ای که شرایط نوری متغیر، بازتابش نور از سطوحی نظیر شیشه عینک (به‌ویژه عینک‌های آفتابی) و اختلال در الگوریتم‌های ردیابی[1]، می‌تواند منجر به بروز خطاهای فاحش در استخراج ویژگی‌های بیومتریک گردد. یافته‌های تجربی نشان می‌دهند که در شرایط محیطی واقعی، نرخ خطای «عدم آشکارسازی»[2] و «آشکارسازی اشتباه»[3] به‌دلیل کاهش فاصله تفکیک‌پذیری میان اجزای چهره افزایش می‌یابد. این ناپایداری فنی در تشخیص جزئیات، به‌ویژه در لحظات کاهش هوشیاری یا تغییرات جزئی در میمیک چهره، اثبات می‌کند که خروجی این سیستم‌ها همواره با درصدی از خطا همراه بوده و نمی‌تواند به‌عنوان یک گزاره قطعی در فرایندهای حساس شناسایی مورد استناد قرار گیرد (سیگاری و همکاران، ۱۳۹۱، صص ۴۷-۴۸).

در تحلیل ماهیت عملکردی سیستم‌های تشخیص هویت، باید به این نکته‌ی کلیدی اشاره داشت که سامانه‌های تشخیص چهره[4]، برخلاف انگاره‌های رایج عمومی، فاقد توانایی ارائه‌ی یک «تطبیق قطعی»[5] به معنای مطلق کلمه هستند. خروجی فنی این سامانه‌ها در واقع فهرستی از «نامزدهای احتمالی»[6] بر اساس ضرایب تشابه است و در نهایت، این «اپراتور انسانی» است که باید بر مبنای قضاوت نهایی، هویت فرد را احراز نماید (Jackson, 2019, p. 15).

یکی از کانون‌های بحران‌زا در این فرایند، استفاده از «عکس‌های کاوشگر»[7] با کیفیت نازل (مانند تصاویر استخراج‌شده از دوربین‌های مداربسته) است. ضابطان قضایی گاه برای جبران این نقیصه‌ی فنی، به «ویرایش‌های افراطی» یا «عادی‌سازی»[8] متوسل می‌شوند؛ اقدامات چالش‌برانگیز نظیر جایگزینی اجزای صورت با عکس‌های آرشیوی یا بازسازی چهره با استفاده از تصاویر افراد مشابه (مانند سلبریتی‌ها) به جای متهم واقعی. این مداخلات انسانی نه‌تنها ضریب اطمینان خروجی الگوریتم را به‌شدت تقلیل می‌دهد، بلکه ماهیت ادله را از یک یافته‌ی مستدل علمی به یک «برآورد حدسی» و غیرقابل اتکا تنزل می‌دهد (Jackson, 2019, p. 16). این نقیصه‌ی فنی زمانی که با مداخلات انسانی همراه می‌شود، مفاهیم حقوقی همچون "اذن" و "رضایت" را نیز تحت‌الشعاع قرار می‌دهد. در تحلیل حقوقی، اذن به معنای تمایل قلبی به تعرضی است که علیه حقوق فرد انجام می‌شود (صابری و اکرمی، ۱۴۰۰، ص ۲۴۰). اما، در سیستم‌های تشخیص چهره، اگر اپراتور انسانی با علم به "ویرایش افراطی" تصویر، آگاهانه خروجی سیستم را بپذیرد، در واقع نوعی اذن به ریسک را صادر کرده است. با این حال، باید توجه داشت که اذن اپراتور به استفاده از داده‌های بی‌کیفیت، به معنای سلب مسئولیت کیفری از وی در صورت بازداشت اشتباه اشخاص نیست؛ چراکه در مسائل عمومی و نظم قضایی، افراد نمی‌توانند چیزی را که شارع و مقنن منع کرده (یعنی تعرض به آزادی بی‌گناه)، با رضایت یا اذن خصوصی خود مباح سازند (صابری و اکرمی، ۱۴۰۰، ص ۲۳۸).

از سوی دیگر، قلمرو اشتباهات و مخاطرات هوش مصنوعی تنها به حوزه‌ی کشف جرم محدود نمی‌ماند؛ چراکه این فناوری در دستان مجرمان سایبری، فرایند شناسایی «نقاط آسیب‌پذیر»[9] در زیرساخت‌های شبکه را به‌صورت خودکار در آورده است. الگوریتم‌های یادگیری ماشین اکنون قادر هستند با سرعتی فراتر از توان تحلیل انسانی، ضعف‌های امنیتی نرم‌افزارها را رصد کرده و از آن‌ها به‌عنوان مدخلی برای حملات سازمان‌یافته بهره‌برداری کنند. فراتر از این، هوش مصنوعی منجر به پیدایش نسل نوین «بدافزارهای سازگارپذیر» شده است؛ بدافزارهایی که با تحلیل

1. tracking
2. false negative
3. false positive
4. FRT
5. match
6. candidate list
7. probe photos
8. normalization
9. vulnerabilities

محیط و تدابیر دفاعی، ساختار عملکردی خود را به‌صورت پویا تغییر می‌دهند تا از رادارهای سیستم‌های تشخیص نفوذ پنهان بمانند. این دگردیسی فنی، ماهیت جرایم را از حملات ایستای سنتی به تهدیداتی «پویا، هوشمند و پیش‌بینی‌ناپذیر» بدل ساخته است (صوفی و صالح‌نژاد، ۱۴۰۴، ص ۲).

۲- ماهیت حقوقی اشتباهات سیستم‌های هوشمند[1]

یکی از غامض‌ترین چالش‌های نوظهور در عرصه‌ی حقوق کیفری، نحوه تبیین و احراز «رابطه سببیت» میان خروجی‌های پردازش‌شده توسط سامانه‌های هوشمند و تصمیم نهایی مقام قضایی یا ضابط است. در این چارچوب، پرسش بنیادین این است که آیا خروجی نادرست یک سیستم، صرفاً یک «مداخله فنی»[2] در فرایند تصمیم‌گیری است یا باید آن را ذیل عنوان «تقصیر فنی»[3] تحلیل کرد؟ (کفانی‌نژاد، ۱۴۰۴، ص ۶). در فقدان دکترین و مقررات صریح، مفاهیمی نظیر «اعتماد معقول» به فناوری و «درجه اتکای حرفه‌ای» کاربر به ماشین، به‌عنوان ستون‌های نظری اصلی برای تحلیل حقوقی این مسئولیت و تعیین سهم تقصیر عامل انسانی در مواجهه با خطای ماشین تلقی می‌شوند (کفانی‌نژاد، ۱۴۰۴، ص ۷).

از منظر استانداردهای ادله اثبات دعوی، فناوری تشخیص چهره هنوز به تراز «قابلیت اتکای علمی»[4] نرسیده است تا بتواند به‌عنوان یک دلیل مستقیم و مستقل در محاکم پذیرفته شود. به همین سبب، در رویه قضایی کشورهای پیشرو نظیر ایالات متحده، این فناوری صرفاً در حد یک «سرنخ تحقیقاتی»[5] معتبر است؛ نه یک دلیل اثباتی قاطع (Jackson, 2019, p. 15). این چالش حقوقی زمانی ابعاد بحرانی به خود می‌گیرد که با پدیده‌ی «سوگیری بیومتریک»[6] مواجه می‌شویم؛ به این معنا که سیستم‌های مذکور در شناسایی گروه‌های خاص (زنان، جوانان و رنگین‌پوستان)، نرخ خطای معنادار و بسیار بالاتری نشان می‌دهند. این عدم دقت ساختاری، ماهیت اشتباهات را از یک «خطای اتفاقی» به یک «نقص سیستماتیک» بدل می‌کند؛ امری که می‌تواند منجر به نقض حقوق بنیادین و بازداشت‌های خودسرانه، بدون وجود «دلیل کافی»[7] گردد (Jackson, 2019, p. 17). این نقایص سیستماتیک ایجاب می‌کند که در تحلیل ماهیت حقوقی این اشتباهات، از ظرفیت‌های "نظارت مشارکتی" (مانند نظارت والدین و ارائه‌دهندگان خدمات) به‌عنوان لایه‌های صیانت‌گر بهره‌برداری شود. استقرار فناوری‌های تأیید هویت که با مشارکت بخش خصوصی و در قالب "حکمرانی فضای سایبر" طراحی می‌شوند، می‌توانند به‌عنوان یک مکمل کنترلی، از اتکای انحصاری و خطرناک نهادهای رسمی به خروجی‌های غیرقطعی هوش مصنوعی بکاهند و مسئولیت شناسایی را میان حاکمیت و جامعه مدنی توزیع نمایند (نظری و همکاران، ۱۴۰۰، ص ۱۵۸).

در قلمرو جرایم مالی نیز هوش مصنوعی با بهینه‌سازی تکنیک‌های «مهندسی اجتماعی»، نرخ موفقیت حملات فیشینگ[8] و فارمینگ[9] را به شکلی تصاعدی افزایش داده است. مهاجمان با تحلیل هوشمند الگوهای رفتاری کاربران در شبکه‌های اجتماعی، پیام‌های جعلی را چنان دقیق و شخصی‌سازی شده طراحی می‌کنند که تمییز دادن آن‌ها برای یک کاربر متعارف، عملاً ناممکن می‌گردد. همچنین، بهره‌گیری از تکنیک‌هایی چون «دوقلوهای شرور»[10] در

1. legal nature of ai errors
2. technical intervention
3. technical fault
4. scientific reliability
5. investigatory lead
6. biometric bias
7. probable cause
8. phishing
9. pharming
10. evil twins

شبکه‌های بی‌سیم و هدایت ترافیک درگاه‌های بانکی به صفحات جعلی از طریق دستکاری سرورهای DNS ، از جمله پیچیدگی‌هایی است که به واسطه ابزارهای هوشمند، با دقت و وسعت عملیاتی مخرب‌تری اجرا می‌شوند (صوفی و صالح‌نژاد، ۱۴۰۴، صص ۱۰-۱۲).

در نهایت، برای مقابله با نگرانی‌های ناشی از نقض حریم خصوصی و مخاطرات برون‌سپاری تصاویر محرمانه به سرورهای ابری، راهکارهای فنی پیشرفته‌ای نظیر ترکیب «مدارهای مبهم»[1] و «رمزنگاری تمام هم‌ریخت»[2] پیشنهاد شده است. در این رویکرد صیانت‌محور، کاربر تصویر خود را به‌صورت رمزنگاری‌شده برای سرور ارسال کرده و سرور بدون دسترسی مستقیم به هویت واقعی تصویر، الگوریتم‌های تشخیص (مانند Eigenfaces) را بر روی داده‌های رمز شده اجرا می‌نماید. این فرایند بدیع تضمین می‌کند که نه تصویر خصوصی کاربر و نه پایگاه‌داده‌ی مرجع سرور برای طرف مقابل فاش نگردد، در حالی‌که خروجی نهایی با دقتی بالا و در محیطی امن حاصل می‌شود (آزاد سنجری و همکاران، ۱۳۹۳، ص ۸۴۲).

۳- مبانی انتساب مسئولیت کیفری[3]

در رهیافتی تطبیقی، دکترین غالب در نظام‌های حقوقی پیشرو، به‌ویژه در ایالات متحده، بر این اصل استوار است که بهره‌گیری از ابزارهای هوشمند به هیچ روی رافع تکلیف مقام انسانی در اعمال قضاوت نهایی نیست. از این‌رو، مسئولیت نهایی حقوقی کماکان بر عهده‌ی «کاربر انسانی» باقی می‌ماند (کفانی‌نژاد، ۱۴۰۴، ص ۸). در مقابل، این مسئولیت صرفاً به کاربر محدود نشده و شرکت‌های توسعه‌دهنده‌ی نرم‌افزار نیز در صورت احراز نقص در طراحی، سوگیری در داده‌آموزی[4] یا کوتاهی در ارائه‌ی هشدارهای لازم نسبت به محدودیت‌های فنی سامانه، واجد مسئولیت خواهند بود (کفانی‌نژاد، ۱۴۰۴، ص ۹). این در حالی است که در نظام حقوقی ایران، فقدان سازوکاری شفاف برای تفکیک مرز میان «خطای انسانی» و «نقص فنی»، نوعی «ابهام تقنینی» را پدید آورده است که پیامد آن می‌تواند ایجاد مسئولیت مضاعف برای کاربر و یا تضییع حقوق دفاعی متهم باشد (کفانی‌نژاد، ۱۴۰۴، ص ۷).

تحلیل مبانی انتساب مسئولیت در قلمرو هوش مصنوعی، در گرو تبیین تلاقی میان «تقصیر انسانی» و «استقلال الگوریتمی»[5] است که چالش‌برانگیزترین حوزه در حقوق کیفری مدرن محسوب می‌شود. در حقوق ایران، مسئولیت کیفری بر بنیان اصل «شخصی بودن مجازات» استوار است و منحصراً بر اشخاصی بار می‌شود که از ارکان «ادراک»، «اراده» و «فاعلیت انسانی» برخوردار باشند (الجیزانی و حیدری، ۱۴۰۴، ص ۴). در تکمیل این چالش، باید به ریشه‌های هستی‌شناختی این بن‌بست توجه داشت؛ چراکه تفاوت ماهوی میان "اراده مبتنی بر نفس" در انسان و "اراده الگوریتمی" در سازه‌های مصنوعی، انتساب مسئولیت را با دشواری مضاعف روبه‌رو می‌کند.

هوش مصنوعی در فلسفه حقوق، نه یک "آبژه" (ابزار محض) و نه یک "سوژه" (انسان کامل) است، بلکه پدیده‌ای "نیمه‌خودمختار" تلقی می‌شود. از این‌رو، تلاش برای تطبیق وضعیت این فناوری با نهادهای سنتی فقهی و حقوقی همچون "عاقله" یا "مسئولیت ناشی از اشیا"، به‌دلیل فقدان آگاهی اخلاقی در ماشین، راهگشا نخواهد بود و اصرار بر تفسیر قوانین فعلی برای مواجهه با این پدیده، تنها به ابهام در دستگاه قضایی منجر می‌شود (زندی و رفیعی علوی، ۱۴۰۳، صص ۹۵-۱۰۲). با این وجود، پیچیدگی کنش‌های هوشمند، ضرورت گذار از نظریات سنتی را ایجاب کرده و نظریه‌های نوینی همچون «مسئولیت توزیعی» و «تقصیر ساختاری» را به متن مباحث حقوقی کشانده است. بر مبنای

---

1. garbled circuits
2. FHE
3. attribution of criminal liability
4. data training
5. algorithmic autonomy

این نظریات، مسئولیت نباید صرفاً بر دوش اپراتور نهایی سنگینی کند، بلکه باید میان زنجیره‌ای از کنشگران شامل طراح، برنامه‌نویس و بهره‌بردار توزیع گردد (رضاپور، ۱۴۰۴، ص ۱). این نظام‌مندسازی مسئولیت در فضای سایبر، مستلزم عبور از الگوهای سنتی مسئولیت فردی و حرکت به سمت "مسئولیت جمعی و همگانی" است.

یافته‌های پژوهشی در حوزه‌ی رسانه‌های الکترونیک نشان می‌دهند که در صورت ارتکاب جرم توسط ابزارهای دیجیتال خودکار، نباید تنها به مباشر بسنده کرد، بلکه بر اساس نظریه "غفلت در نظارت"، مسئولیت کیفری می‌تواند به صورت "نیابتی و اعتباری" گریبان‌گیر مدیران بستر و ارائه‌دهندگان خدمات هوشمند نیز باشد. این رویکرد، با تکیه بر الگوهای مسئولیت در قانون مطبوعات، پاسخی به خلأهای ناشی از دشواری اثبات سوء نیت در الگوریتم‌های پیچیده است و مسئولیت را بر دوش نهادهایی می‌گذارد که کنترل فنی بستر را در اختیار دارند (لطفی، ۱۳۹۸، صص ۴۱۲-۴۱۸).

در این زنجیره‌ی مسئولیت، مفهوم "برائت" نیز نقشی کلیدی ایفا می‌کند. در فقه و حقوق کیفری ایران، اگر پیش از انجام عملی (مانند مداخله فنی در پرونده) از ذی‌نفع اخذ برائت شود، ذمه شخص فارغ می‌گردد؛ مشروط بر این‌که موازین علمی و فنی رعایت شده باشد (صابری و اکرمی، ۱۴۰۰، ص ۲۴۱). اما، چالش اصلی در هوش مصنوعی اینجاست که آیا توسعه‌دهنده می‌تواند با اخذ برائت از نهادهای دولتی، مسئولیت "نقص عضو معنوی" (آبرو و آزادی) ناشی از خطای سیستم را از خود سلب کند؟ پاسخ حقوقی به این پرسش منفی است؛ زیرا برائت تنها زمانی رافع مسئولیت است که منجر به "بی‌احتیاطی و عدم مهارت" نشود. لذا، اخذ برائت توسط شرکت‌های فناوری، نمی‌تواند ابزاری برای توجیه عدم مهارت یا قصور در طراحی الگوریتم‌های حساس جنایی باشد (صابری و اکرمی، ۱۴۰۰، ص ۲۴۱).

در سطحی فراتر، اتحادیه‌ی اروپا با عبور از نگاه ابزاری به فناوری و پذیرش مفهوم «مسئولیت غیرمستقیم»، به سمت شناسایی ماهیت حقوقی جدیدی تحت عنوان «شخصیت الکترونیکی»[1] برای سامانه‌های مستقل گام برداشته است. هدف از این نوآوری حقوقی آن است که در صورت وقوع آسیب، امکان انتساب مسئولیت به موجودیتی فراتر از یک «شیء» ساده فراهم آید و خلأهای ناشی از استقلال ماشین پوشش داده شود (الجیزانی و حیدری، ۱۴۰۴، ص ۸).

تحقق مسئولیت کیفری در حقوق ایران، مستلزم احراز ارکان "عقل، بلوغ و اختیار" است که انتساب مستقیم بزه به هوش مصنوعی را با بن‌بست اهلیت جنایی مواجه می‌سازد. از سوی دیگر، اعمال مجازات بر موجودیت‌های هوشمند با چالش "عدم امکان نیل به اهداف کیفر" روبه‌رو است؛ چراکه هوش مصنوعی فاقد درک از مفاهیمی چون درد، رنج یا ترس است و به تبع آن، کارکرد "بازدارندگی خاص" و "سزادهی عادلانه" در قبال آن منتفی می‌گردد. با این حال، با توجه به استقلال رأی سامانه‌های نوین، ضرورت انقلابی در تفکر حقوقی جهت شناسایی هوش مصنوعی به عنوان "شخص الکترونیکی"[2] گریزناپذیر است تا خلأ ناشی از مصونیت این عوامل در زنجیره‌ی مسئولیت پوشش داده شود؛ هرچند تا زمان استقرار شخصیت حقوقی مستقل، مسئولیت نهایی همچنان بر عهده "فاعل معنوی" شامل برنامه‌نویس یا کاربر انسانی است (کاوه و بارانی، ۱۴۰۳، صص ۵۴-۵۸).

٤- رابطه سببیت و چالش «جعبه سیاه»[3]

بسیاری از سامانه‌های مدرن، به‌ویژه الگوهای مبتنی بر «یادگیری عمیق»[4]، واجد ابهامی ذاتی در فرایند پردازش داده‌ها هستند؛ به‌گونه‌ای که حتی طراحان سیستم نیز قادر به تبیین دقیق و منطقی چرایی صدور یک خروجی خاص نمی‌باشند.

1. electronic personality
2. electronic person
3. causality & black box challenge
4. deep learning

این پدیده که در ادبیات فنی و حقوقی تحت عنوان «جعبه سیاه»[1] شناخته می‌شود، مانعی جدی در مسیر شفافیت قضایی است (کفانی‌نژاد، ۱۴۰۴، ص ۱۰).

در دعاوی مسئولیت، احراز «رابطه سببیت» میان فعل و نتیجه، منوط به تفکیک دقیق نقش‌ها است. لذا، چنانچه ضابط یا مقام قضایی نتواند اثبات کند که تصمیم نهایی صرفاً برآیند تحلیل محض سامانه بوده و یا ناشی از مداخله و قضاوت فردی، اثبات رابطه سببیت با دشواری‌های جدی مواجه خواهد شد. برای برون‌رفت از این معضل، استقرار سامانه‌هایی که واجد قابلیت «ردیابی تصمیم»[2] باشند در پرونده‌های جنایی الزامی است تا در صورت وقوع خطا، منشأ دقیق آن، اعم از نقص ساختاری فناوری یا خطای کاربر، قابل شناسایی باشد (کفانی‌نژاد، ۱۴۰۴، ص ۱۰). در واقع، این «شفافیت الگوریتمی»، نقشی بنیادین در احراز رابطه‌ی سببیت و تفکیک تقصیر انسانی از نواقص فنی ایفا می‌کند (کفانی‌نژاد، ۱۴۰۴، ص ۱۰).

در این چارچوب، پاسخ‌گویی[3] صرفاً به معنای پذیرش مسئولیت نیست، بلکه شامل الزام به "توجیه و تبیین"[4] منطق پشت تصمیمات ماشین است. تجارب قضایی در ایالات متحده (مانند پرونده ویلی آلن لینچ) نشان می‌دهد که ناتوانی ضابطان در درک نحوه‌ی عملکرد رتبه‌بندی الگوریتم، منجر به بازداشت‌های ناعادلانه شده است؛ در این پرونده، پلیس حتی از تبیین مقیاس ضریب تشابهی که منجر به شناسایی متهم شده بود، عاجز ماند. این پدیده مؤید آن است که بدون "قابلیت تفسیر"[5]، خروجی‌های هوش مصنوعی نمی‌توانند مبنای مشروعی برای سلب آزادی اشخاص باشند (Almeida et al., 2022, p. 385).

فراتر از ابهامات فنی، اثبات رابطه سببیت در این حوزه با پدیده‌ی مخاطره‌آمیز «خلأ سببی»[6] روبه‌رو است؛ وضعیتی که در آن به سبب ماهیت «خودآموز»[7] سیستم، اتصال فعل مجرمانه به یک فاعل انسانی مشخص عملاً ناممکن یا بسیار دشوار می‌گردد (الجیزانی و حیدری، ۱۴۰۴، ص ۱۴). از آنجا که بسیاری از الگوریتم‌های یادگیری عمیق در قالب یک جعبه سیاه عمل کرده و رفتاری پیش‌بینی‌ناپذیر حتی برای طراحان خود دارند، مفهوم سنتی «قصد کیفری»[8] دیگر توان درک و تحلیل افعالی را ندارد که از این سامانه‌های مستقل ناشی می‌شوند (الجیزانی و حیدری، ۱۴۰۴، ص ۱۴). برای حل این چالش در نظام‌های دادرسی مدرن، تأکید بر دو اصل «شفافیت» و «قابلیت تفسیر»[9] گریزناپذیر است. به عبارت دیگر، منطق تصمیم‌گیری هوش مصنوعی باید به‌گونه‌ای ردیابی و تحلیل شود که امکان تفکیک میان «خطای فنی غیرقابل پیش‌بینی» و «سهل‌انگاری انسانی» فراهم گردد (رضاپور، ۱۴۰۴، ص ۷).

۵- تأثیر بر اصول دادرسی منصفانه[10]

به‌کارگیری سامانه‌های تشخیص چهره در فرایندهای کیفری، مرزهای سنتی میان «اماره قضایی» و «دلیل قطعی» را با چالشی بنیادین مواجه کرده است. از آنجا که تصمیم‌گیری بر مبنای خروجی‌های «جعبه سیاه» واجد ابهامات ساختاری است، تحول در نظام‌های ادله اثبات دعوی ایجاب می‌کند که به‌تدریج ادله علمی و یافته‌های آزمایشگاهی، جایگزین نظام ادله معنوی سنتی گردند؛ زیرا برخلاف شواهد ذهنی مانند اقرار و شهادت که به‌شدت در معرض خطا، فشار روانی و تصرفات انسانی قرار دارند، دلایل مادی و عینی حاصل از فرایندهای فنی به‌دلیل ماهیت غیرقابل‌انکارشان، از ضریب

1. black box
2. decision traceability
3. accountability
4. justification
5. explain ability
6. causal gap
7. self-learning
8. mens rea
9. explain ability
10. impact on fair trial

اطمینان و اعتبار قضایی بالاتری برخوردارند. استناد به این یافته‌های علمی، ضمن تضمین دستیابی به حقیقت، نقش مؤثری در تقلیل اشتباهات قضایی و احقاق حقوق بنیادین اشخاص ایفا کرده و در صورت تعارض با ادله سنتی، می‌تواند صحت آن‌ها را مورد تردید جدی قرار دهد (بابایی و همکاران، ۱۳۹۷، ص ۱۴۳). استناد بلاشرط و بدون تحقیق به این داده‌ها می‌تواند منجر به تضییع حق دفاع متهم گردد (کفانی‌نژاد، ۱۴۰۴، ص ۹).

علاوه بر این، در فرایند کارشناسی، چنانچه متخصصان پلیس علمی صرفاً بر تحلیل بصری تصویر تمرکز کرده و از ماهیت عملکردی و پیچیدگی‌های هوش مصنوعی غافل بمانند، احتمال تحلیل نادرست منشأ خطا افزایش می‌یابد. از این‌رو، حضور متخصصان فناوری اطلاعات در هیئت‌های کارشناسی جهت «صحت‌سنجی مستقل الگوریتم‌ها»، یکی از الزامات گریزناپذیر برای استقرار عدالت در پرونده‌های متکی بر فناوری‌های نوین است (کفانی‌نژاد، ۱۴۰۴، ص ۱۱). تأثیر این اشتباهات بر "حق حیات و مصونیت از تعرض" که از مهم‌ترین حقوق بنیادین هستند، شدیدترین واکنش کیفری را ایجاب می‌کند (صابری و اکرمی، ۱۴۰۰، ص ۲۴۴).

در دعاوی مرتبط با هوش مصنوعی، رضایت یا اعتماد اولیه ضابطان به سامانه، نباید به‌عنوان یک "عامل موجهه" تلقی شود که وصف مجرمانه را از بازداشت‌های غیرقانونی بزداید. از منظر حقوق عمومی، از آنجا که مجازات‌ها و قواعد دادرسی برای حفظ مصالح عالیه (دین، عقل، جان و مال) وضع شده‌اند، هرگونه کوتاهی در صحت‌سنجی ادله‌ی دیجیتال که منجر به نقض این مصالح شود، تحت هیچ عنوان با تمسک به "اذن اداری" یا "رضایت سازمانی" قابل توجیه نیست (صابری و اکرمی، ۱۴۰۰، ص ۲۵۰).

یکی از ابعاد حیاتی که در تحلیل خدمات هوش مصنوعی مغفول مانده، مفهوم «قابلیت وارسی»[1] است. در پروتکل‌های امنیتی نوین، همچون VGHFR، نه‌تنها حریم خصوصی کاربران صیانت می‌شود، بلکه نهاد قضایی قادر است صحت محاسبات انجام‌شده توسط سرور را به‌طور دقیق رصد کند. به بیانی دیگر، سرور موظف است در کنار نتیجه‌ی تشخیص چهره، یک «اثبات محاسباتی» نیز ارائه دهد تا فرضیه‌ی تقلب فنی یا خطای عمدی در شناسایی افراد که مستقیماً به بازداشت‌های اشتباه منجر می‌شود، منتفی گردد. این پروتکل با بهینه‌سازی پیچیدگی زمانی، امکان وارسی درستی الگوریتم را حتی برای نهادهایی با قدرت پردازشی محدود فراهم می‌آورد (آزاد سنجری و همکاران، ۱۳۹۳، ص ۸۴۴).

از سوی دیگر، «سوگیری‌های الگوریتمی»[2] پتانسیل ایجاد رفتارهای تبعیض‌آمیز ساختاری علیه گروه‌های خاص (زنان، جوانان و اقلیت‌ها) را دارا هستند که این امر در تعارض آشکار با اصول بنیادین حقوق بشر و «اصل برائت» قرار دارد (الجیزانی و حیدری، ۱۴۰۴، ص ۱۱). در نظام حقوقی ایران، به سبب فقدان رویه‌ی قضایی تخصصی و ضعف در آموزش‌های فنی ضابطان و قضات، خطر «اعتماد کورکورانه به ماشین»[3] به‌شدت احساس می‌شود. این پدیده می‌تواند حق دفاع متهم را با چالشی جدی روبه‌رو سازد (رضاپور، ۱۴۰۴، ص ۷). لذا، صیانت از حقوق بنیادین ایجاب می‌کند که هرگونه تشخیص ماشین، صرفاً به‌عنوان یک «قرینه»[4] تلقی شده و لزوماً تحت «نظارت و بازبینی مستقیم انسانی»[5] قرار گیرد (رضاپور، ۱۴۰۴، ص ۲).

در این راستا، گذار از نگاه صرفاً فنی به رویکردی "جامعه‌محور" در قالب سیاست جنایی مشارکتی ضرورت می‌یابد؛ بدین معنا که برای جلوگیری از "اعتماد کورکورانه"، باید از ظرفیت سازمان‌های مردم‌نهاد و انجمن‌های سواد رسانه‌ای جهت ارتقای دانش دیجیتالی کنشگران عدالت کیفری بهره جست. در واقع، پیشگیری کنشی از اشتباهات

1. verifiability
2. algorithmic bias
3. automation bias
4. clue
5. human oversight

سیستمی، بیش از آن‌که بر محدودیت‌های فنی استوار باشد، نیازمند "آموزش مستمر" و درونی‌سازی هنجارهای اخلاقی در مواجهه با ابزارهای هوشمند است تا از تبدیل شدن این فناوری‌ها به ابزار "تلقین پنهان" به شهود و قضات جلوگیری شود (نظری و همکاران، ۱۴۰۰، ص ۱۷۲).

علاوه بر سوگیری‌های داده‌ای، باید به "سوگیری طراحان" نیز توجه داشت؛ از آنجا که بخش بزرگی از توسعه‌دهندگان فناوری در غرب را مردان سفیدپوست تشکیل می‌دهند، ناخودآگاه "سوگیری ورودی"[1] در الگوریتم‌ها نهادینه می‌شود که نتیجه‌ی آن، خطای فاحش در شناسایی زنان و رنگین‌پوستان است. برای مقابله با این بی‌عدالتی ساختاری، انجام "ارزیابی تأثیر حقوق بشر"[2] در کنار "ارزیابی تأثیر حفاظت از داده‌ها"[3] پیش از استقرار هرگونه سامانه تشخیص چهره توسط پلیس، یک ضرورت اخلاقی و قانونی است تا از تبدیل شدن هوش مصنوعی به ابزاری برای تبعیض نژادی یا جنسیتی جلوگیری شود (Almeida et al., 2022, p. 379).

در نهایت، پنهان‌کاری در استفاده از این فناوری، تهدیدی جدی علیه دادرسی منصفانه محسوب می‌شود. در بسیاری از پرونده‌ها، پلیس از تشخیص چهره برای شناسایی اولیه بهره می‌برد، اما در مراحل بعدی با تمسک به سایر ادله، نقش کلیدی الگوریتم را مخفی نگاه می‌دارد (Jackson, 2019, p. 17). این رویکرد، حق متهم برای چالش کشیدن اعتبار ادله را سلب می‌کند. افزون بر این، مواجهه شهود با متهمی که پیش‌تر توسط ماشین انتخاب شده، نوعی «تلقین پنهان»[4] ایجاد می‌کند؛ چراکه شاهد تحت تأثیر این باور قرار می‌گیرد که سیستم هوشمند پیشاپیش امکان خطا را منتفی کرده است، در حالی‌که ممکن است سیستم صرفاً «شبیه‌ترین فرد» به مجرم را برگزیده باشد، نه لزوماً خود مجرم را (Jackson, 2019, p. 19).

گذر از پلیس "واکنشی" به پلیس "پیش‌گیرانه" از طریق استقرار فناوری‌های تشخیص چهره، اگرچه ضرورتی فنی در مقابله با بزهکاری مدرن است، اما نباید منجر به "شیفتگی تکنولوژیک" و نادیده انگاشتن اصول راهبردی دادرسی گردد. استقرار این سامانه‌ها بدون وجود "قانون خاص" و "تنظیم‌گری شفاف"، مخاطراتی جدی نظیر نظامی شدن عدالت کیفری، نقض "فرض برائت" و ایجاد سیاست جنایی توتالیتر را به همراه دارد. بر این اساس، رعایت پنج اصل کلیدی منشور اخلاقی اروپا شامل: "احترام به حقوق بنیادین"، "کیفیت و امنیت داده‌ها"، "عدم تبعیض"، "شفافیت الگوریتمی" و "سلطه کاربر انسانی بر ماشین"، پیش‌شرط مشروعیت حقوقی استفاده از هوش مصنوعی در فرایند تعقیب کیفری است تا از تبدیل شدن ابزارهای قضایی به وسایل نقض حریم خصوصی در فضاهای عمومی جلوگیری شود (شیخوند و همکاران، ۱۴۰۲، صص ۸۸-۹۶).

٦- چالش‌های جرم‌انگاری و خلأهای قانونی در ایران

تحلیل ساختار قانونی ایران نشان می‌دهد که برخی مفاهیم مبهم مانند "ضرورت" در ماده ۵۰ قانون آیین دادرسی کیفری، به ضابطان "اختیارات بدون تحدید" می‌دهد که می‌تواند منجر به برخوردهای سلیقه‌ای در استفاده از ابزارهای هوشمند شود (دارایی و اختری، ۱۴۰۱، ص ۱۶۵). علاوه بر این، "اجراناپذیری اداری" برخی مواد قانونی، مانند الزام به اعلام مراتب تحت‌نظر بودن به دادستان ظرف یک ساعت، در کنار کمبود نیروی انسانی و تجهیزات، پلیس را در موقعیتی قرار می‌دهد که برای رعایت ظواهر قانونی، به روش‌های میان‌بر تکنولوژیک متوسل شود. لذا، پیشگیری از نقض حقوق شهروندی مستلزم "اصلاح فنی قوانین جرم‌زا" و شفاف‌سازی اختیارات ضابطان در مواجهه با ادله‌ی دیجیتال است (دارایی و اختری، ۱۴۰۱، ص ۱۶۶).

---

1. input bias
2. HRIA
3. DPIA
4. suggestiveness

بغرنج‌ترین چالش پیش‌روی نظام قضایی ایران، خصیصه‌ی «نوپدید بودن» جرایم مبتنی بر هوش مصنوعی است که در بسیاری از لایه‌ها، فراتر از قلمرو پیش‌بینی قوانین موضوعه (به‌ویژه قانون جرایم رایانه‌ای مصوب ۱۳۸۸) قرار می‌گیرند. این «خلأ تقنینی» در تلاقی با ویژگی‌های ذاتی فضای مجازی نظیر «گمنامی پیشرفته» و «سهولت دسترسی»، فرایند پیچیده‌ی تعقیب، شناسایی و اثبات جرم را با دشواری‌های جدی مواجه ساخته است. به عنوان نمونه، در حوزه جرایم علیه عفت و اخلاق عمومی، هوش مصنوعی با تسهیل تولید هرزه‌نگاری‌های مصنوعی[1] و خلق هویت‌های جعلی، مرزهای سنتی اخلاق و عفت را جابه‌جا نموده است؛ این در حالی است که چارچوب‌های قانونی موجود، هنوز میان جرایم جنسی در اتمسفر فیزیکی و سایبری، به‌ویژه در موارد حساس بزه‌دیدگی کودکان، تفکیکی دقیق، فنی و بازدارنده قائل نشده‌اند (صوفی و صالح‌نژاد، ۱۴۰۴، صص ۸-۹).

علاوه بر خلأهای تقنینی، تمرکز مفرط سیاست جنایی ایران بر "تدابیر وضعی حاکمیتی" (مانند پالایش و فیلترینگ) و غفلت از "پیشگیری مشارکتی"، منجر به تضعیف نظارت‌های هوشمند شده است. این رویکرد سخت‌افزاری، نه‌تنها دقت سیستم‌های تشخیص هویت را در تفکیک محتوای مجرمانه از غیرمجرمانه کاهش می‌دهد، بلکه با ایجاد اشتباه در فیلترینگ، حقوق مشروع کاربران را نیز سلب می‌نماید. لذا، کارآمدی این سامانه‌ها در گرو تفویض بخشی از اختیارات نظارتی به نهادهای مدنی و کاربران است تا با افزایش "دقت و هوشمندی جمعی"، ضریب خطای سیستم‌های تشخیص خودکار تقلیل یابد (نظری و همکاران، ۱۴۰۰، ص ۱۵۳).

در بُعد زیرساختی و برای درک چرایی نفوذ گسترده‌ی این سامانه‌ها در نهادهای امنیتی، باید به کارآمدی الگوریتم‌هایی نظیر Eigenfaces اشاره کرد که بر پایه‌ی «تحلیل مؤلفه‌های اصلی»[2] استوار هستند. این روش به عنوان یکی از ستون‌های بنیادین تشخیص چهره در محیط‌های حساس شناخته می‌شود؛ چراکه در آن، تصاویر به فضایی با ابعاد کاهش یافته (فضای چهره) منتقل می‌شوند تا فرایند محاسبات امن با شتاب بیشتری صورت پذیرد.

یافته‌های حاصل از پیاده‌سازی‌های عملیاتی با ابزارهایی مانند SCAPI مبین آن است که از طریق بهینه‌سازی «تعداد گیت‌های منطقی»[3] می‌توان توازن دقیقی میان امنیت و سرعت برقرار کرد. این بهینه‌سازی اجازه می‌دهد که علی‌رغم اعمال لایه‌های رمزنگاری سنگین (جهت صیانت از داده‌ها)، زمان پاسخ‌گویی در پایگاه‌داده‌های کلان در مقیاس میلی‌ثانیه حفظ شده و کارایی عملیاتی سیستم برای ضابطان قضایی تضمین گردد (آزاد سنجری و همکاران، ۱۳۹۳، ص ۸۴۵). با این حال، سرعت پردازش نباید به بهای قربانی کردن "حریم خصوصی در طراحی" تمام شود.

برخلاف رویکرد متکثر و فاقد متولی واحد در ایالات متحده، الگوی اتحادیه اروپا (مبتنی بر GDPR) بر این اصل استوار است که حفاظت از داده‌ها باید به‌صورت "پیش‌فرض" در معماری فنی سیستم لحاظ گردد. در ایران نیز برای عبور از خلأهای قانونی، پذیرش "اثر بروکسل"[4]، یعنی انطباق با استانداردهای عالی حفاظتی بین‌المللی، می‌تواند راهگشا باشد. این امر مستلزم آن است که شرکت‌های خصوصی تأمین‌کننده‌ی فناوری برای پلیس، مکلف به انتشار عمومی گزارش‌های ارزیابی تأثیر خود باشند تا امکان نظارت عمومی بر امنیت و دقت الگوریتم‌ها فراهم گردد (Almeida et al., 2022, p. 381).

علاوه بر الزامات فنی، در ساحت جبران خسارت نیز می‌توان از ظرفیت‌های موجود در نظام حقوقی ایران بهره جست. در موارد ابهام میان "تقصیر" و "اشتباه" ناشی از به‌کارگیری فناوری‌های نوین قضایی، اصل ۱۷۱ قانون اساسی و ماده ۱۳ قانون مجازات اسلامی، مبنای مستحکمی برای "مسئولیت جبرانی دولت" فراهم می‌آورند. با نگاهی تطبیقی،

1. deep fakes
2. PCA
3. Non-XOR
4. Brussels effect

می‌توان استدلال کرد که هرگاه خسارت، ناشی از "نقص در ابزارهای نظام عدالت" (مانند خطای الگوریتمی در سامانه‌های تشخیص هویت) باشد، مسئولیت از شخص قاضی یا ضابط سلب و بر عهده‌ی دولت قرار می‌گیرد. این رویکرد که بر اصل "حق دادنی است، نه ستاندنی" استوار است، ایجاب می‌کند که دولت به‌جای ارجاع بزه‌دیده به مسیرهای دشوار اثبات تقصیر شرکت‌های فناوری، رأساً اقدام به جبران خسارت نموده و سپس در صورت احراز تقصیر، به عامل اصلی (طراح یا اپراتور) رجوع نماید (رحماندوست، ۱۳۹۹، صص ۶-۱۱).

نتیجه‌گیری و پیشنهادات

پژوهش حاضر نشان داد که ورود سیستم‌های تشخیص چهره به فرایندهای جنایی، اگرچه سرعت کشف جرم را ارتقا بخشیده، اما «عدالت کیفری» را با چالشی از جنس «شبح در ماشین» مواجه ساخته است. یافته‌ها مؤید آن هستند که این فناوری، برخلاف تصور عمومی، یک «دلیل قطعی» نیست، بلکه صرفاً فهرستی از احتمالات را ارائه می‌دهد که در صورت فقدان «شفافیت الگوریتمی» و «نظارت انسانی»، می‌تواند به ابزاری برای نقض سیستماتیک حقوق بنیادین و بازداشت‌های خودسرانه بدل شود.

در تحلیل مسئولیت، مشاهده شد که مدل‌های سنتی «مسئولیت فردی» در برابر «جعبه سیاه» هوش مصنوعی رنگ باخته‌اند. از سویی، بن‌بست اهلیت جنایی هوش مصنوعی و فقدان درک کیفر در ماشین، انتساب مستقیم بزه به الگوریتم را ناممکن می‌سازد و از سوی دیگر، تمرکز صرف بر تقصیر اپراتور انسانی، ناعادلانه و نادیده انگاشتن «نقص‌های طراحی» است. لذا، گذار به سمت «مسئولیت توزیعی» و شناسایی هوش مصنوعی به‌عنوان یک «شخصیت الکترونیکی» در آینده، ضرورتی انکارناپذیر برای صیانت از حریم خصوصی و نظم قضایی می‌باشد.

بر مبنای یافته‌های این تحقیق، پیشنهادات ذیل جهت اصلاح سیاست جنایی و رویه قضایی ارائه می‌گردد:

۱. در سطح تقنینی

- تدوین قانون خاص هوش مصنوعی: وضع مقرراتی جامع جهت تفکیک «هوش مصنوعی ابزاری» از «سیستم‌های خودمختار» و پذیرش «رهیافت ریسک‌محور» به‌جای جرم‌انگاری سنتی.
- اصلاح قانون جرایم رایانه‌ای: گنجاندن مصادیق نوین بزهکاری (نظیر جعل عمیق و سوگیری الگوریتمی) و تعیین الزامات «حفاظت از داده‌ها در طراحی» برای شرکت‌های طرف قرارداد با نهادهای انتظامی.

۲. در سطح قضایی و اجرایی

- تغییر منزلت اثباتی: خروجی سیستم‌های تشخیص چهره در محاکم نباید به‌عنوان «دلیل مستقل» پذیرفته شود، بلکه باید صرفاً منزلت «اماره قضایی» یا «سرنخ تحقیقاتی» داشته باشد که صحت آن مستلزم تأیید توسط هیئت‌های کارشناسی دوگانه (پلیس علمی و متخصصان IT) است.
- فعال‌سازی ظرفیت اصل ۱۷۱ قانون اساسی: در مواردی که به‌دلیل پیچیدگی فنی، مقصر واحد قابل شناسایی نیست، دولت باید بر اساس مسئولیت جبرانی، رأساً خسارات مادی و معنوی بزه‌دیدگان را جبران نماید.

۳. در سطح فنی و اخلاقی

- الزام به شفافیت و وارسی: استقرار پروتکل‌های فنی که امکان «ردیابی تصمیم» ماشین را فراهم آورده و از «اعتماد کورکورانه» ضابطان به خروجی‌های هوشمند جلوگیری کند.
- آموزش و سواد دیجیتال: برگزاری دوره‌های تخصصی برای قضات و ضابطان جهت درک «سوگیری‌های بیومتریک» و محدودیت‌های ذاتی فناوری، تا از تبدیل شدن هوش مصنوعی به یک «تلقین‌گر پنهان» در پرونده‌های حساس جنایی ممانعت به عمل آید.

در نهایت، باید گفت، تکنولوژی نباید به بهای قربانی کردن «اصل برائت» در محراب «سرعت و دقت» ستایش شود؛ چراکه عدالت، حتی اگر با قدم‌های آهسته‌ی انسانی اجرا شود، بسی والاتر از بی‌عدالتی سازمان‌یافته‌ای است که با سرعت میلی‌ثانیه‌ای یک ماشین رخ می‌دهد.

منابع

آزاد سنجری، ف.، عادلی، م.، و برومند، س. (۱۳۹۳). تشخیص چهره امن با استفاده از مدارهای مبهم و تحلیل مؤلفه‌های اصلی. *دومین همایش ملی فناوری اطلاعات، ارتباطات و محاسبات*، کرمان.

الجیزانی، ع. و حیدری، ا. (۱۴۰۴). چالش‌های مسئولیت کیفری در جرایم ناشی از هوش مصنوعی. *فصلنامه دیدگاه‌های حقوق قضایی*.

بابایی، م.، قورچی‌بیگی، م.، و شاکری، ع. (۱۳۹۷). تحول در نظام ادله اثبات دعوای کیفری در پرتو یافته‌های علمی. *مجله مطالعات حقوقی دانشگاه شیراز*، ۱۰(۱)، ۱۱۷-۱۵۴.

دارایی، م.، و اختری، م. (۱۴۰۱). چالش‌های ضابطان دادگستری در مواجهه با حقوق شهروندی در نظام دادرسی ایران. *فصلنامه مطالعات حقوق عمومی*، ۵۲(۱)، ۱۶۱-۱۸۳.

رحماندوست، ج. (۱۳۹۹). مسئولیت مدنی و کیفری دولت در جبران خسارات ناشی از اشتباهات قضایی بر مبنای اصل ۱۷۱ قانون اساسی. *پژوهشنامه حقوق کیفری*، ۱۱(۲)، ۵-۲۵.

رضاپور، م. (۱۴۰۴). تحلیل مبانی مسئولیت توزیع‌شده در سامانه‌های خودمختار هوشمند. *مجله حقوقی دادگستری*.

زندی، م.، و رفیعی علوی، س. م. (۱۴۰۳). مبانی فلسفی-حقوقی اراده در هوش مصنوعی و بازخوانی مفهوم اهلیت جنایی. *فصلنامه نقد رای*، ۱۳(۱)، ۹۰-۱۱۰.

سیگاری، م. ح.، پوررضا، م. ر.، و صباغیان، ن. (۱۳۹۱). بازشناسی چهره در شرایط محیطی متغیر و تحلیل خطاهای بیومتریک. *نشریه علمی پردازش علائم و داده‌ها*، ۹(۱)، ۴۳-۶۰.

شیخوند، م.، صبوری، ع.، و قربانی، ا. (۱۴۰۲). اخلاق و حکمرانی هوش مصنوعی در سیاست جنایی اروپا. *فصلنامه پژوهش حقوق کیفری*، ۱۱(۴۲)، ۸۵-۱۱۲.

صابری، ع.، و اکرمی، م. (۱۴۰۰). تحلیل فقهی-حقوقی مفهوم اذن و برائت در مسئولیت‌های ناشی از فناوری‌های نوین. *دوفصلنامه فقه و حقوق اسلامی*، ۱۲(۲۴)، ۲۳۵-۲۵۵.

صوفی، م.، و صالح‌نژاد، ح. (۱۴۰۴). نسل نوپدید جرایم سایبری در بستر هوش مصنوعی: تهدیدها و چالش‌های قانونی. *فصلنامه انتظام اجتماعی*.

کاوه، س.، و بارانی، م. (۱۴۰۳). هوش مصنوعی و ضرورت شناسایی شخصیت الکترونیکی در نظام حقوقی ایران. *مطالعات حقوقی مدرن*، ۵(۱)، ۵۰-۷۲.

کفانی‌نژاد، م. (۱۴۰۴). *مسئولیت کیفری ضابطان در به‌کارگیری سامانه‌های هوشمند تشخیص هویت*. تهران: انتشارات میزان.

لطفی، و. (۱۳۹۸). مسئولیت کیفری مدیران بستر و ارائه‌دهندگان خدمات در رسانه‌های الکترونیک. *فصلنامه مطالعات حقوق کیفری و جرم‌شناسی*، ۶(۲)، ۴۱۱-۴۳۰.

نظری، م.، عباسی، ا.، و صادقی، ح. (۱۴۰۰). سیاست جنایی مشارکتی در حکمرانی فضای مجازی و پیشگیری از جرایم هوشمند. *مجله حقوقی دادگستری*، ۸۵(۱۱۴)، ۱۵۱-۱۷۵.

Almeida, D., Shmarko, K., & Loke, E. (2022). Algorithmic bias and the Brussels effect: Regulating AI in criminal justice. *International Journal of Law and Information Technology*, *30*(4), 375-395.

Jackson, K. (2019). *The faulty logic of facial recognition: A report on biometric errors and civil liberties*. Washington, DC: Georgetown Law Center on Privacy & Technology.

استناد به این مقاله: افروزی، هدیه، و محمودی، امیررضا. (۱۴۰۵). مسئولیت کیفری ناشی از اشتباهات سیستم‌های تشخیص چهره در پرونده‌های جنایی. *فصلنامه تحقیقات نوین میان‌رشته‌ای حقوق*، ۶(۲)، ۱۰۵-۱۱۹.