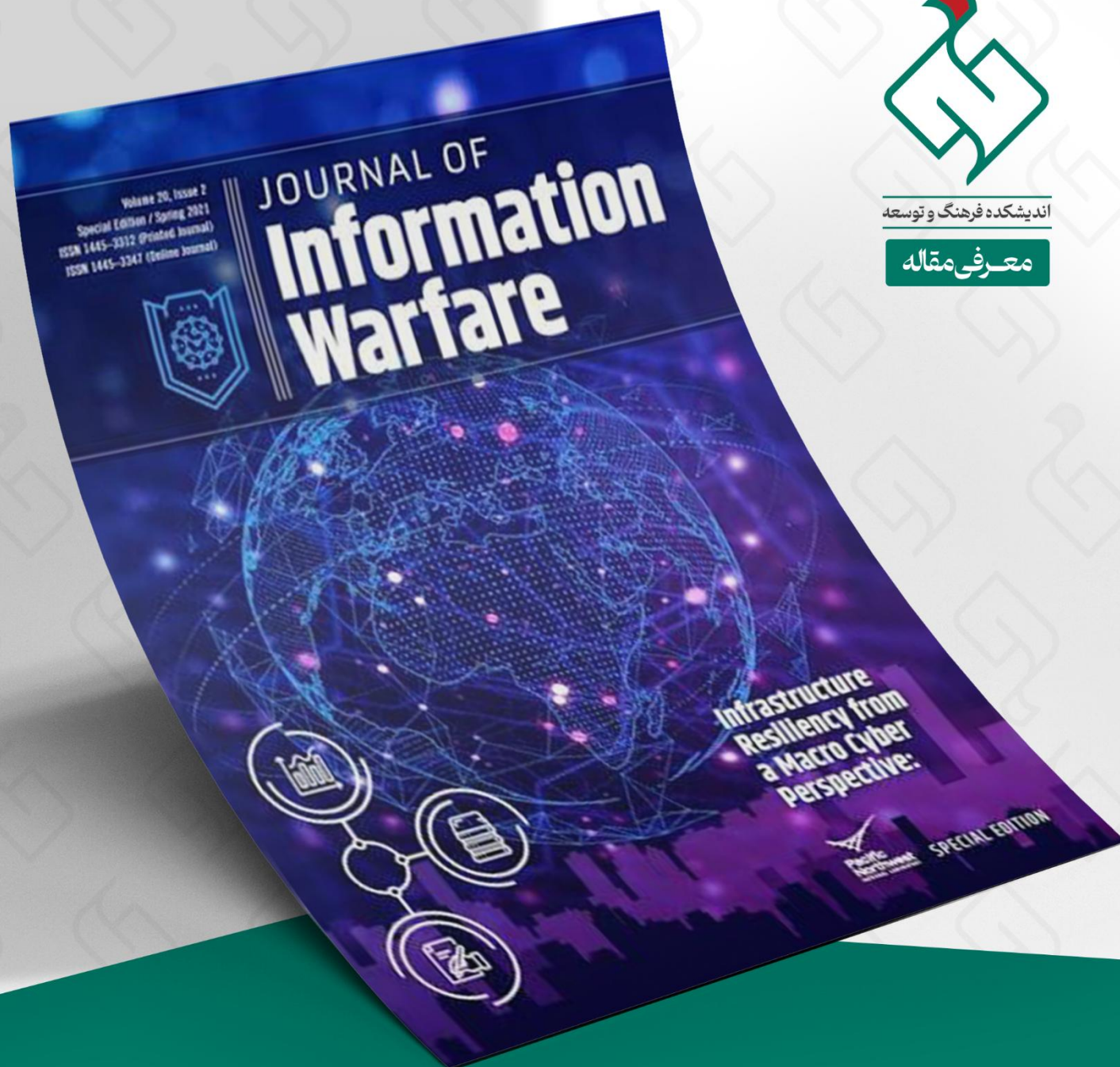




اندیشکده فرهنگ و توسعه

معرفی مقاله



فراتر از دیپفیک

تصاویر متحرک مصنوعی و آینده تاریخ

نویسنده: دبلیو آرکی | فصلنامه جنگ اطلاعات | ۲۰۲۴



چکیده:

این مقاله به بررسی نقش ابزارهای هوش مصنوعی مولد (AI) در تولید تصاویر متحرک مصنوعی می‌پردازد - به طور خاص، چگونگی استفاده از این تصاویر در کمپین‌های آنلاین انتشار اطلاعات نادرست و تأثیر عمیق آنها بر آرشیوهای فیلم تاریخی. محتوای دستکاری‌شده توسط هوش مصنوعی، به ویژه تصاویر متحرک، تأثیری بسیار فراتر از محیط فعلی جنگ اطلاعاتی (IW) خواهد داشت و با پیامدهای ناشناخته، به حوزه‌های ناشناخته آرشیوهای تاریخی بصری نفوذ خواهد کرد. این مقاله همچنین سناریوهای جنگ اطلاعاتی را که در آنها انواع جدیدی از کمپین‌های بلندمدت انتشار اطلاعات نادرست ممکن است ظهور کنند، بررسی می‌کند و در پایان به استراتژی‌های بالقوه تأیید و مهار می‌پردازد.

کلمات کلیدی: دیپ‌فیک، تصاویر متحرک مصنوعی، رسانه مصنوعی، هوش مصنوعی، اطلاعات گمراه‌کننده، جنگ اطلاعاتی، آرشیوهای تاریخی، فیلم مستند



خلاصه مبسوط مقاله:

این مقاله نقش ابزارهای هوش مصنوعی مولد را در تولید تصاویر متحرک مصنوعی بررسی می‌کند و بر دو پیامد اصلی آن تمرکز دارد: استفاده از این تصاویر در کارزارهای اطلاعات نادرست و تأثیر آن‌ها بر آرشیوهای تاریخی تصویری. نویسنده استدلال می‌کند که گسترش این فناوری، اعتبار مستند تصاویر تاریخی را با چالشی اساسی مواجه می‌سازد؛ زیرا تصاویری که پیش‌تر به‌عنوان شواهدی از وقوع رویدادها پذیرفته می‌شدند، اکنون می‌توانند به‌آسانی تولید یا دستکاری شوند. مقاله برای تبیین مسئله، میان اطلاعات نادرست، اطلاعات درست با کاربرد فریبکارانه و اطلاعات کذب با قصد فریب تمایز می‌گذارد. در مفهوم نخست، فرد ممکن است بدون قصد فریب، محتوایی نادرست را منتشر کند؛ در مفهوم دوم، اطلاعات واقعی با زمینه‌ای گمراه‌کننده بازنشر می‌شود؛ و در مفهوم سوم، انتشار اطلاعات نادرست با هدف فریب صورت می‌گیرد. نویسنده همچنین جنگ اطلاعاتی را کشمکش بر سر فرایندهای اطلاعات و ارتباطات معرفی می‌کند که در محیطی متشکل از زیرساخت فیزیکی، شبکه‌ها و سامانه‌های اطلاعاتی و لایه شناختی فعالیت می‌کند. تمرکز مقاله بر لایه شناختی و تأثیر تصاویر مصنوعی بر ادراک تاریخ، هویت و واقعیت است.

مقاله در بخش بررسی وضعیت فناوری هوش مصنوعی و تصاویر متحرک مصنوعی، روند تحول این فناوری را از پیشرفت‌های یادگیری ماشین و شبکه‌های عصبی تا ظهور مدل‌های مولد تشریح می‌کند. نویسنده توضیح می‌دهد که شبکه‌های مولد تخصصی از دو شبکه عصبی تشکیل شده‌اند که یکی تصاویر را تولید می‌کند و دیگری آن‌ها را ارزیابی می‌کند تا خروجی به هدف موردنظر نزدیک شود. رمزگذارهای خودکار تغییرپذیر نیز از دیگر رویکردهای تولید تصویر معرفی می‌شوند. افزایش حجم داده‌های آموزشی، توان پردازشی و قدرت مدل‌ها، امکان تولید تصاویر واقع‌گرایانه‌تر را فراهم کرده است. مقاله در ادامه، میان جلوه‌های ویژه رایانه‌ای متعارف و ابزارهای جدید هوش مصنوعی تمایز می‌گذارد و بر کاهش هزینه و دشواری تولید محتوای مصنوعی تأکید می‌کند. توسعه ابزارهای تولید ویدئو از متن، از جمله مدل سورا که در زمان نگارش مقاله معرفی شده بود، نشان می‌دهد که تولید ویدئوهای باکیفیت و منسجم به‌سرعت در حال پیشرفت است. نویسنده همچنین به فناوری تولید و شبیه‌سازی صدا اشاره می‌کند که امکان تقلید از صدای افراد و تولید گفتار دلخواه را فراهم می‌سازد. ترکیب این قابلیت‌ها، تولید ویدئوهای مصنوعی همراه با صدا را به امکانی گسترده‌تر و در دسترس‌تر تبدیل می‌کند.

در بخش مربوط به هوش مصنوعی و اطلاعات نادرست، مقاله نشان می‌دهد که دستکاری اطلاعات و تصاویر پدیده‌ای تازه نیست، اما فناوری دیجیتال و شبکه‌های اجتماعی مقیاس، سرعت و دسترسی به آن را تغییر داده‌اند. در گذشته، جعل تصاویر به صحنه‌سازی یا دستکاری در فرایند عکاسی نیاز داشت؛ سپس نرم‌افزارهایی مانند فتوشاپ امکان تغییر تصاویر را با هزینه کمتر فراهم کردند. با گسترش شبکه‌های اجتماعی، اطلاعات گمراه‌کننده می‌توانند پیش از انتشار اصلاحیه‌ها و بررسی‌های صحت‌سنجی به مخاطبان گسترده‌ای برسند. الگوریتم‌هایی که تعامل و کلیک را در اولویت قرار می‌دهند، انتشار محتوای هیجانی و افراطی را تشویق می‌کنند و حباب‌های اطلاعاتی نیز می‌توانند دیدگاه‌های موجود را تقویت کنند. مقاله برای توضیح نمودهای این وضعیت به گسترش دیپ‌فیک‌ها، تصاویر جعلی مرتبط با شخصیت‌های سیاسی و تصاویر مصنوعی مرتبط با جنگ‌ها اشاره می‌کند. نمونه‌هایی مانند ویدئوی جعلی تسلیم رئیس‌جمهور اوکراین و تصاویر تولیدشده از آتش‌سوزی ساختمان پنتاگون، ظرفیت این فناوری برای ایجاد سردرگمی را نشان می‌دهند. نویسنده تأکید می‌کند که تصاویر به دلیل قدرت تأثیرگذاری عاطفی و شناختی خود، در شکل‌دهی به برداشت



افراد از جهان و رویدادهای تاریخی نقش مهمی دارند. از این رو، گسترش تصاویر مصنوعی می‌تواند اعتماد به ارزش مستند تصاویر خبری و تاریخی را تضعیف کند.

در بخش آینده تصاویر تاریخی، مقاله به وضعیت آرشیوهای تصویری و آسیب‌پذیری آن‌ها در برابر محتوای مصنوعی می‌پردازد. از زمان پیدایش تصاویر متحرک در اواخر قرن نوزدهم، حجم گسترده‌ای از فیلم‌ها و ویدئوها برای ثبت رویدادهای روزمره، دستاوردهای انسانی، جنگ‌ها، نسل‌کشی‌ها و دیگر وقایع تاریخی تولید شده است. این مواد در آرشیوهای دولتی، دانشگاهی و تجاری نگهداری می‌شوند و برای تولید اخبار، مستندها و سایر محصولات تصویری مورد استفاده قرار می‌گیرند. با این حال، آرشیوهای جهان از استانداردهای مشترک و یکپارچه‌ای برخوردار نیستند و محدودیت منابع، تعدد قالب‌های قدیمی و دشواری دیجیتالی‌سازی، حفاظت از این مجموعه‌ها را پیچیده می‌کند. دیجیتالی‌سازی، امکان دسترسی و استفاده مجدد از تصاویر را افزایش می‌دهد، اما هم‌زمان آن‌ها را از حامل‌های فیزیکی اولیه جدا می‌کند و زمینه دستکاری و بازتولیدشان را فراهم می‌سازد. نویسندگان هشدار می‌دهد که تصاویر تاریخی موجود می‌توانند با افزودن یا تغییر جزئیات، معنایی متفاوت پیدا کنند و حتی نسخه‌های مصنوعی از رویدادهای گذشته تولید شوند. سپس این تصاویر ممکن است دوباره وارد آرشیوهای تجاری یا نهادی شوند. در چنین وضعیتی، مرز میان اسناد واقعی و تصاویر بازسازی‌شده مبهم می‌شود و امکان استفاده از آرشیوها برای بازنویسی تاریخ افزایش می‌یابد.

مقاله در بخش پیامدها و کارزارهای بلندمدت اطلاعات نادرست، دو نگرانی مرتبط را مطرح می‌کند. نگرانی نخست، امکان تولید تصاویر تاریخی مصنوعی برای استفاده در فیلم‌های مستند و روایت‌های رسانه‌ای است. این فناوری می‌تواند هزینه دسترسی به تصاویر آرشیوی را کاهش دهد و امکان بازنمایی تجربه جوامع به حاشیه‌رانده‌شده‌ای را فراهم کند که رویدادهای مربوط به آن‌ها هرگز ثبت نشده است. با این حال، استفاده از تصاویر مصنوعی ممکن است اعتبار تصاویر مستند واقعی را نیز تضعیف کند. نگرانی دوم، دشواری تشخیص محتوای مصنوعی و ضرورت احراز اصالت اسناد تاریخی موجود است. نویسندگان سپس سناریوهایی را بررسی می‌کند که در آن‌ها تصاویر جعلی برای تحریک منازعات قومی، اجتماعی و سیاسی یا بازنمایی دروغین بی‌عدالتی‌های تاریخی به کار می‌روند. انتشار فیلم‌هایی که تصاویر واقعی و جعلی را ترکیب می‌کنند، می‌تواند خشم و میل به انتقام را برانگیزد و از بررسی دقیق رسانه‌ها پیشی بگیرد. همچنین ممکن است تصاویر تازه‌ای از رویدادهای بحث‌برانگیز تولید و در اینترنت یا آرشیوها منتشر شوند و پس از دریافت اعتبار ظاهری، به منابعی برای تولید محتوای بعدی تبدیل گردند. مقاله این امکان را در چارچوب جنگ اطلاعاتی شناختی بررسی می‌کند و به ظرفیت چنین فناوری‌هایی برای تقویت راهبردهای چین، روسیه و دیگر بازیگران دولتی و غیردولتی در ایجاد سردرگمی، تضعیف اعتماد نهادی و تثبیت روایت‌های تاریخی تحریف‌شده اشاره دارد.

در بخش اخلاق هوش مصنوعی، مهار و راستی‌آزمایی، نویسندگان مجموعه‌ای از رویکردهای فنی و نهادی را بررسی می‌کند. نخستین رویکرد، هدایت توسعه هوش مصنوعی با چارچوب‌های اخلاقی، ممیزی‌های معتبر و سازوکارهای پاسخ‌گویی است. مقاله به تلاش‌هایی برای ایجاد استانداردهای اخلاقی و نظارتی، از جمله فعالیت‌های سازمان یونسکو و ابتکارهای مرتبط با پاسخ‌گویی الگوریتمی، اشاره می‌کند. در سطح فنی، ابزارهای تشخیص تصاویر تولیدشده با هوش مصنوعی و روش‌های واترمارک‌گذاری دیجیتال مطرح می‌شوند. سامانه SynthID می‌تواند نشانه‌ای نامرئی برای مخاطب عادی در پیکسل‌های تصویر تعبیه کند تا منشأ مصنوعی آن قابل شناسایی باشد. ابزار PhotoGuard نیز با تغییر پیکسل‌ها می‌کوشد امکان دستکاری تصاویر به وسیله الگوریتم‌های هوش مصنوعی را محدود کند. مقاله همچنین استاندارد منشأ و اصالت محتوا، موسوم به C2PA، را توضیح می‌دهد که هدف آن ثبت فراداده‌های تغییرناپذیر



درباره منشأ تصویر و انتقال اطلاعات مربوط به ویرایش‌های بعدی است. ثبت رمزنگاری‌شده فراداده‌ها و استفاده از فناوری بلاک‌چین نیز از راهکارهای پیشنهادی برای حفظ قابلیت ردیابی اسناد محسوب می‌شوند. با وجود این، نویسنده تأکید می‌کند که هیچ‌یک از این ابزارها راه‌حلی کامل نیستند. کیفیت پایین تصاویر، ضعف تعمیم‌پذیری سامانه‌های تشخیص، دشواری پذیرش گسترده استانداردها، امکان جعل نشانه‌های اصالت و دور زدن سازوکارهای ثبت منشأ، اثربخشی این روش‌ها را محدود می‌کنند.

در ادامه همین بحث و در نتیجه‌گیری مقاله، نویسنده بر ضرورت حفاظت پیشگیرانه از آرشیوهای تاریخی تأکید می‌کند. به اعتقاد او، یکی از اقدامات اولیه باید بررسی وضعیت آرشیوهای دولتی و تجاری و شناخت سازوکارهایی باشد که این مؤسسات برای حفاظت از اصالت تصاویر خود به کار می‌گیرند. در کنار توسعه ابزارهای فنی، لازم است تصاویر تاریخی مهم اولویت‌بندی شوند، منشأ آن‌ها تا حد امکان به حامل اصلی بازگردانده شود و اطلاعات مربوط به اصالتشان به شیوه‌ای قابل اعتماد ثبت گردد. مقاله به فعالیت آزمایشگاه استارلینگ اشاره می‌کند که از رمزنگاری و مفاهیم وب ۳ برای ثبت، نگهداری و راستی‌آزمایی مواد اولیه استفاده می‌کند. ایجاد برچسب رسمی اصالت یا بهره‌گیری از ممیزی شخص ثالث می‌تواند امکان تشخیص تصاویر تأییدشده را برای مستندسازان، تولیدکنندگان محتوای آرشیوی و دیگر استفاده‌کنندگان فراهم سازد. نویسنده در عین حال، هزینه، زمان و فضای ذخیره‌سازی موردنیاز برای حفاظت از حجم گسترده اسناد دیجیتال‌نشده را از موانع این فرایند می‌داند. در جمع‌بندی، مقاله استدلال می‌کند که گسترش تصاویر مصنوعی، پذیرش خودکار تصویر به‌عنوان مدرک مستند را از میان می‌برد و پیامدهایی برای سیاست، صنعت، جامعه و فرایندهای حقوقی دارد. آموزش عمومی درباره توانایی‌های هوش مصنوعی، تقویت سواد رسانه‌ای انتقادی و تدوین پروتکل‌های پایدار برای احراز اصالت تصاویر تاریخی، از اقدامات ضروری معرفی می‌شوند. نویسنده همچنین احتمال رقابت مستمر میان تولیدکنندگان تصاویر مصنوعی و توسعه‌دهندگان ابزارهای تشخیص را مطرح می‌کند و بر ضرورت آمادگی جوامع برای مواجهه با این وضعیت تأکید دارد.

- **موضوع اصلی:** بررسی پیامدهای تصاویر متحرک تولیدشده با هوش مصنوعی مولد برای جنگ اطلاعاتی، عملیات فریب و اعتبار اسناد تصویری تاریخ، با تأکید بر آسیب‌پذیری آرشیوهای تاریخی در برابر جعل و بازنویسی گذشته.
- **مسئله محوری:** گسترش ابزارهای تولید تصاویر و ویدئوهای مصنوعی، اعتبار ضمنی تصاویر تاریخی را به چالش می‌کشد. تصویری که پیش‌تر به‌عنوان شواهدی از رویدادهای واقعی پذیرفته می‌شدند، ممکن است دستکاری شوند یا با نسخه‌هایی مصنوعی جایگزین شوند که روایتی متفاوت از گذشته ارائه می‌کنند.
- **چارچوب تحلیلی:** پیوند میان هوش مصنوعی مولد، اطلاعات نادرست، اطلاعات گمراه‌کننده و جنگ اطلاعاتی، با تمرکز بر لایه شناختی محیط اطلاعاتی؛ یعنی جایی که دستکاری تصاویر می‌تواند بر ادراک، عواطف، حافظه جمعی و برداشت افراد از واقعیت اثر بگذارد.
- **یافته کلیدی:** دیجیتال‌سازی آرشیوهای تصویری، در کنار تسهیل دسترسی و استفاده مجدد از اسناد، امکان بازتولید و دستکاری آن‌ها را نیز افزایش می‌دهد. در نبود استانداردهای مشترک برای احراز اصالت، تصاویر مصنوعی یا دستکاری‌شده ممکن است وارد آرشیوهای تجاری و دولتی شوند و در مقام اسناد تاریخی مورد استفاده قرار گیرند.
- **سناریوهای آینده:** مقاله احتمال شکل‌گیری کارزارهای بلندمدت فریب را بررسی می‌کند؛ از شبکه‌های اجتماعی متشکل از هویت‌های مصنوعی تا تولید تصاویر جعلی از رویدادهای تاریخی برای تشدید تنش‌های قومی، اجتماعی و سیاسی. چنین



تصاویری می‌توانند با برانگیختن خشم و احساس بی‌عدالتی، به ابزار تأثیرگذاری بر افکار عمومی و بی‌ثبات‌سازی اجتماعی تبدیل شوند.

- **راهکارهای پیشنهادی:** توسعه روش‌های تشخیص محتوای تولیدشده با هوش مصنوعی، ایجاد استانداردهای معتبر و قابل دسترس برای احراز اصالت تصاویر آرشیوی، ثبت منشأ و سابقه تغییرات محتوا و طراحی سازوکارهای حفاظتی برای جلوگیری از ورود اسناد جعلی به مخازن تاریخی.
- **استدلال اصلی:** مسئله دیپ‌فیک صرفاً تولید تصاویر جعلی یا فریب کاربران در شبکه‌های اجتماعی نیست؛ خطر عمیق‌تر آن است که مرز میان سند واقعی و بازنمایی مصنوعی گذشته تضعیف شود. حفاظت از تاریخ تصویری به سازوکارهای فنی، رویه‌های نهادی و استانداردهای احراز اصالت نیاز دارد تا امکان تشخیص و حفظ اسناد معتبر فراهم بماند.



اندیشکده فرهنگ و توسعه

باشگاه مقالات فرتو