PISSN : 2571-9904 - EISSN : 2602-7763

# المخاوف الأخلاقية من الاستخدامات السلبية لتقنيات الذكاء الاصطناعي: تقنية التزييف العميق أنموذجا

# *Ethical Concerns about the Negative Uses of AI Technologies: Deepfakes Technology as a model*

د. الأسد صالح الأسد (مخبر الدراسات التاريخية والأثرية).

المركز الجامعي مرسلي عبد الله – تيبازة .(alasad1990@yahoo.fr).

**تاريخ الاستلام**: 01/ 04/ 2022 **تاريخ القبول**: 21/ 05 / 2022 **تاريخ النشر**: 30/ 06/ 2022

**ملخص:**

أصبح الذكاء الاصطناعي بتقنياته المتعددة، يُوظف في مختلف مناحي حياتنا اليومية، فمن الهواتف الذكية التي نستخدمها يومياً، إلى السيارات ذاتية القيادة، والرجال الآليين (الروبوتات) وإلى الأسلحة وغيرها من التقنيات الذكية. ولا يختلف اثنان حول أهمية هذا العلم الحديث وفوائده الكثيرة. لكن ما قد يخُتلف حوله، هو الاستخدامات السلبية لعدد من تقنيات الذكاء الاصطناعي، والتهديدات المختلفة التي يمكن أن تحدثها هذه التقنيات على حياة الإنسان ومستقبله.

وهو ما نهدف إلى توضيحه من خلال هذه المقالة العلمية بشكل جلي، مع تسليط الضوء على تقنية "التزييف العميق" الحديثة، كأنموذج لتلك التقنيات التي أصبحت تستخدم بشكل سلبي ولأغراض "شريرة"، سببت العديد من المشاكل للأفراد والحكومات في كثير من دول العالم المختلفة.

**الكلمات الدلالية**: الاستخدامات السلبية؛ التزييف العميق؛ الذكاء الاصطناعي؛ المخاوف الأخلاقية

**Abstract**:

Artificial intelligence, with its multiple technologies, has become employed in various aspects of our daily lives, from the smart phones we use daily, to self-driving cars, robots, weapons, and other smart technologies. No two disagree about the importance of this modern science and its many benefits. But what may differ about it, is the negative uses of artificial intelligence technologies, and the various threats that this technology can pose to human life and future.

That's what we will seek through this article to clarify clearly, highlighting the modern "deepfakes" technology, as a model for those technologies, which are being used negatively and for evil purposes, they have caused problems for individuals & governments in many different countries of the world.

**Keywords:** *Artificial Intelligence; Deepfakes Technology; Ethical Concerns; Negative Uses*

## مقدمة:

يعد مجال الذكاء الاصطناعي مجالاً واسعاً ومثيراً للجدل على المستوى العالمي، فمنذ نشأة هذا العلم خلال خمسينيات القرن الماضي وحتى يومنا هذا، لا تكاد النقاشات العلمية والجدل ينتهي حول تقنياته المتعددة والمتنوعة، والتي سهلت من مهام الحياة اليومية بشكل كبير خاصة في الدول المتقدمة، إلا أن هناك جانب آخر مهم، يمكن القول بأنه الجانب "المظلم" من جوانب هذا العلم الحديث نسبياً، وهو التقنيات الذكية الموجهة للاستخدامات ذات الأهداف السلبية أو الشريرة.

هذا الجانب الأخير، هو ما نريد أن نسلّط عليه الضوء في هذا المقال، بغية التعرف أولاً على هذا العلم الحديث باختصار غير مخل، ثم التعريج على أهم المخاوف الأخلاقية للاستخدامات السلبية لتقنيات علم الذكاء الاصطناعي، مع التركيز على تقنية "التزييف العميق" والتي سنفرد لها محور خاص في هذه البحث، نظراً لاختيارنا لها كأنموذج ممثل لتقنيات الذكاء الاصطناعي الأكثر إثارة للجدل مؤخراً، محاولين تنوير القارئ حول هذا الموضوع من ناحية، ومن ناحية أخرى تقديم الحلول أو الاقتراحات التي يمكن أن تسهم في الحد من تأثيراته السلبية على المستخدمين لتقنياته المختلفة، انطلاقاً مما توصل إليه الباحثون والمختصون في هذا المجال الحيوي.

وانطلاقاً مما تقدم، فإن إشكالية هذا المقال، تتمحور حول السؤال التالي: إلى أي مدى يمكن أن يشكّل الذكاء الاصطناعي بتقنياته المختلفة تهديداً للإنسان؟

كما نطرح التساؤلات الفرعية التالية:

- ما هي أهم المخاوف الأخلاقية من الاستخدامات السلبية لتقنيات الذكاء الاصطناعي؟
- كيف يمكن الحد من تهديدات تقنيات الذكاء الاصطناعي بشكل عام؟
- ماهي تقنية "التزييف العميق" (Deepfakes)، وما أهم تأثيراتها السلبية على الفرد والمجتمع؟
- ما أهم الحلول التي يمكن من خلالها مواجهة التأثيرات السلبية لتقنية "التزييف العميق"؟

## مفهوم الذكاء الاصطناعي (Artificial Intelligence):

يتكون مفهوم الذكاء الاصطناعي من كلمتين، الأولى: الذكاء (Intelligence) وتعني القدرة على الفهم أو التفكير، والثانية: الاصطناعي (Artificial)، وتشير إلى شيء مصنوع أو غير طبيعي (موسى وبلال. 2019.ص 18). وقد كان "جون مكارثي" أحد رواد منظمة العفو الدولية، أول من حدد مصطلح الذكاء الاصطناعي عام 1955، على النحو التالي: "الهدف من الذكاء الاصطناعي هو تطوير آلات تتصرف وكأنها ذكية" (المرجع نفسه. ص 18).

كما قدم كثير من الباحثين والكتاب العديد من التعريفات لمفهوم الذكاء الاصطناعي، لعل من أبرزها، تعريف (بلاي ويتباي- Blay Whitby[i])، حيث يعرفه بأنه: "دراسة للسلوك الذكي (في البشر والحيوانات والآلات)، كما أنه يمثل محاولة لإيجاد السبل التي يمكن بها إدخال مثل هذا السلوك على الآلات الاصطناعية" (ويتباي. 2008. ص 15). ويعرفه (مارفن لي مينسكي-Marvin Lee Minsky) بأنه: بناء برامج الكمبيوتر التي تنخرط في المهام التي يتم إنجازها بشكل مرضٍ من قبل البشر، وذلك لأنها تتطلب عمليات

عقلية عالية المستوى مثل التعلّم الإدراكي وتنظيم الذاكرة والتفكير النقدي (موسى وبلال. مرجع سابق. ص 20).

ويعرفه الباحثان عبد الله موسى وأحمد حبيب بلال، بأنه: "طريقة لصنع حاسوب، أو روبوت يتم التحكم فيه بواسطة الكمبيوتر، أو برنامج يفكر بذكاء، بنفس الطريقة التي يفكر بها البشر الأذكياء" (المرجع السابق. ص 20). ويضيف الباحثان تعريفاً آخر للذكاء الاصطناعي، على أنه: "علم صنع الآلات التي تقوم بأشياء تتطلب ذكاء إذا قام بها الإنسان" (المرجع نفسه).

ويمكن أن نعرف الذكاء الاصطناعي بأنه: "*العلم الذي يدرس السلوكات الذكية عند الكائنات الحية، ويوظفها من خلال برمجيات حاسوبية متطورة، لأداء مهام ووظائف ذكية تحاكي تلك التي يقوم بها الإنسان خاصة، وذلك بهدف تبسيط حياة الإنسان اليومية ولأغراض أخرى متعددة*".

## 2ـ أنواع الذكاء الاصطناعي:

يمكن تقسيم أنواع الذكاء الاصطناعي إلى ثلاثة أنواع رئيسة، تتراوح من رد الفعل البسيط إلى الإدراك والتفاعل، وذلك على النحو التالي (شادي عبد الوهاب وآخرون. 2018. ص 2):

**2-1- الذكاء الاصطناعي الضيّق أو الضعيف (Narrow AI or Weak AI)**: وهو أبسط أشكال الذكاء الاصطناعي، وتتم برمجة الذكاء الاصطناعي للقيام بوظائف معينة داخل بيئة محددة، ويعتبر تصرفه بمنزلة رد فعل على موقف معيّن، ولا يمكن له العمل إلا في ظروف البيئة الخاصة به، ومن الأمثلة على ذلك الروبوت "ديب بلو"، الذي صنعته شركة (IBM)، والذي هزم "جاري كاسباروف" بطل الشطرنج العالمي.

**2-2- الذكاء الاصطناعي القوي أو العام (General AI or Strong AI)**: يتميّز على جمع المعلومات وتحليلها، وعمل تراكم خبرات من المواقف التي يكتسبها، والتي تؤهله لأن يتخذ قرارات مستقلة وذاتية. ومن الأمثلة على ذلك، السيارات ذاتية القيادة، وروبوتات الدردشة الفورية، وبرامج المساعدة الذاتية الشخصية.

2-3- الذكاء الاصطناعي الخارق (Super AI): وهي نماذج لا تزال تحت التجربة بالقدرة وتسعى لمحاكاة الإنسان، ويمكن هنا التمييز بين نمطين أساسيين، الأول: يحاول فهم الأفكار البشرية، والانفعالات التي تؤثر على سلوك البشر، ويملك قدرة محدودة على التفاعل الاجتماعي. أما الثاني، فهو أنموذج لنظرية العقل، حيث تستطيع هذه النماذج التعبير عن حالتها الداخلية، وأن تتنبأ بمشاعر الآخرين ومواقفهم وتتفاعل معها، فهي الجيل القادم من الآلات فائقة الذكاء.

## 3ـ أهم المخاوف الأخلاقية من الاستخدامات السلبية لتقنيات الذكاء الاصطناعي:

سنتناول في هذا المحور، أهم ثلاثة موضوعات تثير قلق الباحثين والمراقبين في مجال تقنيات الذكاء الاصطناعي، والاستخدامات السلبية لها في الحاضر والمستقبل، وهذه الموضوعات هي: أتمتة الوظائف، وبرامج مراقبة المدنيين، والأسلحة ذاتية التشغيل.

**3-1- إحلال الأتمتة والذكاء الاصطناعي مكان العمّال والموظفين:**

شاهدنا منذ عدة أشهر كيف قدمت لنا الصين أحدث مذيعيها لنشرة الأخبار على القناة الحكومية الدولية الصينية، وكان عبارة عن رجل آلي متقن الصنع تسيّره برمجيات ذكاء اصطناعي، لا يختلف في ملامحه ولا في صوته عن المذيعين من بني البشر، الأمر الذي طرح تساؤلات لعل أبرزها، هل ستقضي هذه الآلات الذكية على مستقبل الكثير من وظائف بني البشر؟!

فمع ظهور الأتمتة [ii] واسعة الانتشار والذكاء الاصطناعي، يبرز التوظيف باعتباره مجالاً آخر من المجالات التي تستعد لخوض اضطرابات كبيرة. والخوف المشترك هنا، هو أن تحل الأتمتة والذكاء الاصطناعي مكان العمال البشر في سوق العمل، ما سيؤدي إلى ارتفاع البطالة الجامحة (أوسويا ويسلر الرابع. 2017. ص 2).

**3-2- استخدام الحكومات لبرامج الذكاء الاصطناعي لمراقبة المدنيين:**

من الأمثلة العميقة على هذه المخاطر نشر الحكومات أدوات اصطناعية لمراقبة المدنيين (...) تعكس أعمال مراقبة الحكومة في أفضل حالاتها نية الحكومة بالعمل، إلا أن النيّة قد لا تحمل الثقل المعنوي أو القانوني نفسه كالأعمال الفعلية. لكنه قد يصعب إقناع الآخرين بهذا الفرق عندما تكون الحكومة قمعية. (...) ويمكن للمراقبة غير العادلة مهما كانت تتحلى بوجاهة قانونية، أن تؤدي دور أداة ترسّخ عدم المساواة. وتمكن زيادة تطور الأدوات الاصطناعية كل الحكومات المتشددة – القمعيّة والخيّرة منها على حد سواء – من ممارسة المراقبة (المرجع نفسه. ص 7).

وتقدم الصين أوضح مثال في هذا الشأن، حيث أعدت الحكومة الصينية برنامج مراقبة يعتمد على الذكاء الاصطناعي، مزود بكاميرات مراقبة منتشرة في مختلف المدن والقرى الصينية، بإمكان هذا البرنامج ومن خلال تلك الكاميرات مراقبة تحركات المواطنين ومعرفة أفعالهم وتحركاتهم، حيث يوجد لكل مواطن صورة مخزنة بخوادم البيانات، يعمل ذلك البرنامج على ربط الصور المرسلة من خلاله، بملف الشخص المراقب، ليتم إظهار معلوماته الشخصية كاملة عند الاشتباه به أو لقيامه بعمل ما لا يروق للجهات الحكومية المراقِبة.

**3-3- الأسلحة ذاتية التشغيل:**

يتمثل أحد أهم تطبيقات الذكاء الاصطناعي العسكرية في ظهور "الأسلحة ذاتية التشغيل" (Autonomous Weapon System)، والتي تعرف بأنها "أي نظام تسليحي يتمتع بالاستقلالية في القيام بوظائفه الحيوية، أي أنه يستطيع اتخاذ قرارات تتعلق بالقيام بالبحث والرصد وتحديد وتعقب واختيار ومهاجمة الأهداف من دون تدخل من البشر"، وذلك وفقاً لتعريف اللجنة الدولية للصليب الأحمر (شادي عبد الوهاب وآخرون. مرجع سابق. ص 3).

وتستطيع الأسلة ذاتية التشغيل، والتي تعرف أيضاً باسم "الروبوتات المستقلة الفتاكة" أو "الروبوتات القاتلة"، البحث عن الأهداف وتحديدها ومهاجمتها، بما في ذلك البشر، من دون تدخل أي إنسان في توجيهها !! (المرجع نفسه. ص 3).

## 4- تقنية التزييف العميق (The Deepfakes Technology):

انتشرت خلال الأشهر الماضية عدد من مقاطع الفيديو على مواقع التواصل الاجتماعي لشخصيات شهيرة، على غرار الرئيس الأمريكي الأسبق "باراك أوباما" وهو يتحدث عن الرئيس الأمريكي "دونالد ترامب" وينتقده، ورئيسة مجلس النواب الأمريكي "نانسي بيلوسي" وهي تتعاطى الخمر وتهذي بشكل غريب، والنجم السينمائي الأمريكي "توم كروز" وهو يلعب بشكل سخيف، ... وغيرهم من أولئك المشاهير. لكن اتضح فيما بعد، أن مقاطع الفيديو تلك كانت مزيفة، تم إنتاجها بواسطة الذكاء الاصطناعي، بالاعتماد على تقنية "التزييف العميق" أو الـ (Deepfakes Technology). فماهي هذه التقنية المعتمدة على الذكاء الاصطناعي، وماهي أضرارها ومخاطرها وتهديداتها؟

### 4-1- مفهوم "التزييف العميق" (The Deepfakes):

من الناحية القانونية يعرّف "التزييف العميق" (Deepfakes) على أنه: مقطع فيديو تم إنشاؤه بقصد الخداع، ويبدو أنه يصور شخصاً حقيقياً يقوم بفعل لم يحدث في الواقع (مجلس جودة الحياة الرقمية. 2021. ص 4). كما يمكن تعريف "التزييف العميق" (Deepfakes) على أنه محتوى مرأي أو صوتي أو كلاهما تم التلاعب به باستخدام الذكاء الاصطناعي وتقنية برمجيات متقدمة لتزييف حقيقة الأفراد والأشياء والأماكن والأحداث. ويبدو هذا المحتوى المزيف قريباً من الواقع، وقد يجد عامة الناس صعوبة في اكتشافه (المرجع نفسه. ص 4).

كما يُعرّف مكتب مساءلة الحكومة الأمريكية (GAO)، التزييف العميق بأنه:" أي تسجيل فيديو، أو صورة، أو صوت يبدو حقيقياً، ولكنه تم التلاعب به بواسطة الذكاء الاصطناعي" (Persons. 2020. P.1).

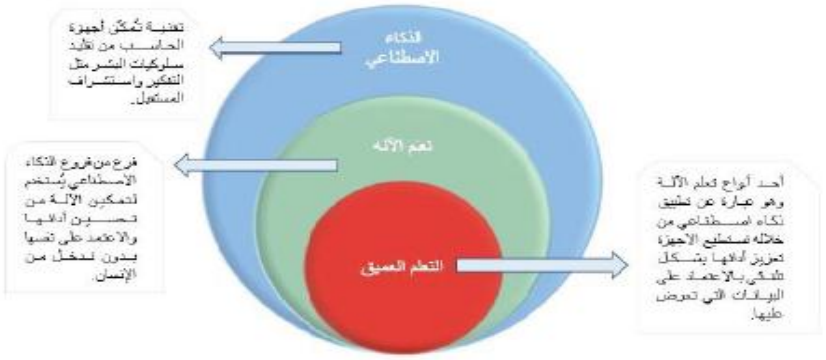
وتعد هذه التقنية واحدة من التقنيات الحديثة المعتمدة على التعلم العميق والتعلم الآلي، طورها "أيان غودفالو" وأطلق عليها اسم (Deepfakes)، أي التزوير العميق (قاسم. 2020. ص 12). ومن الممكن لهذه التكنولوجيا أن تكون خطيرة جداً، حيث تتيح التلاعب بمقاطع الفيديو، وتسهل تلفيق أقوال وأفعال لشخص ما دون علم منه (المرجع السابق. ص 12).

فالتزييف العميق عبارة عن أدوات قوية يمكن أن تستخدم للاستغلال والتضليل. فيمكن من خلال التزييف العميق التأثير في الانتخابات، والتقليل من الثقة، ولكن حتى الآن تم استخدامها بشكل رئيس في إنتاج المواد الإباحية غير المتفق عليها (Persons. op cit. P.1).

**2-4- أنواع التزييف العميق:**

تقنية التزييف العميق، تعتمد على ما يسمى "بالتعلّم العميق"، والذي يمثل فرع من فروع تعلّم الآلة [iii]، الذي تعتمد فكرته على تقليد عمل الخلايا العصبية الموجودة في العقل البشري من خلال ابتكار شبكة عصبية اصطناعية (Artificial Neural Network) تستطيع تحليل كميات ضخمة من البيانات غير المنظمة مثل اللغات المختلفة والصور وترجمتها عبر تمريرها من خلال الشبكة العصبية للتعرف عليها عبر عدة مراحل، ومن هنا جاء مصطلح "العميق" (مجدي. 2020. ص 7). وتشمل تطبيقات التعلم العميق على سبيل المثال: التعرف على الكلام والأصوات والصور (المرجع نفسه. ص 7).

**الشكل 1: يوضح العلاقة بين الذكاء الاصطناعي وتعلّم الآلة والتعلم العميق**



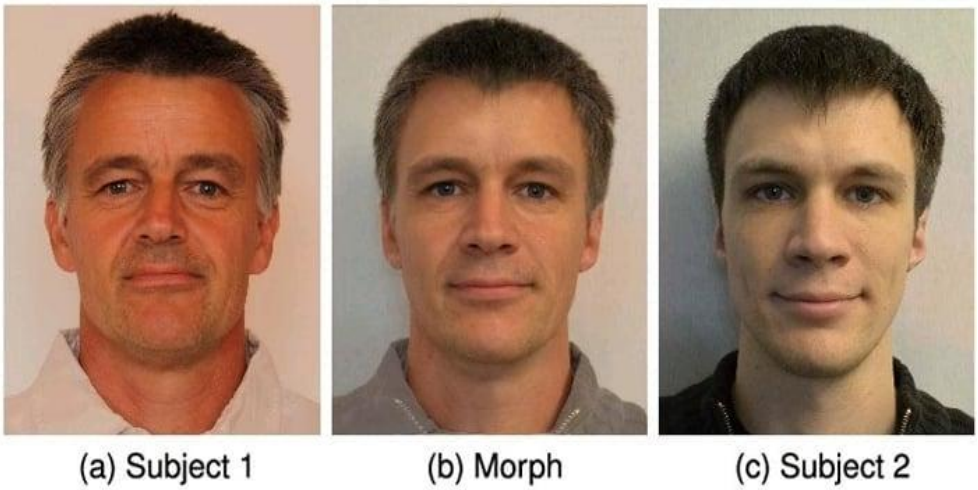

**المصدر: (مجدي. 2020. ص7).**

وترتبط الاستخدامات الأكثر شيوعاً لتقنية "التزييف العميق" (Deepfakes) بما يلي (مجلس جودة الحياة الرقمية. مرجع سابق. ص 10):

**أولاً- المحتوى المرئي:** ويقصد به، استخدام تقنية التزييف العميق في إنشاء الصور ومقاطع الفيديو. ويتم ذلك من خلال:

أ- تبديل الوجه باستخدام خوارزميات التشفير وفك التشفير (Encoder/Decoder Algorithms) لتركيب الخريطة الرقمية (Digital Map) لوجه شخص معين على وجه شخص آخر. حيث تستخدم خوارزمية التشفير آلاف الصور لدراسة ملامح الوجه لدى شخصين مختلفين، ثم تكتشف أوجه التشابه بينهما وتختصرها إلى ميزات مشتركة وتضغط الصور. بعد ذلك يتم تدريب خوارزمية ذكاء اصطناعي ثانية، تسمى بخوارزمية فك التشفير على كيفية استعادة الوجوه من الصور المضغوطة. وبما أن الوجهين مختلفان، تتم برمجة الخوارزمية الأولى لاستعادة وجه الشخص الثاني، ولتبديل الوجهين يتم تزويد معلومات فك التشفير (Decoder Algorithm) ببيانات الصور المشفرة الخاصة بالوجه الآخر. والشكل رقم:2 يوضح هذه العملية.

**الشكل رقم 2: عملية تشكيل صورة جديدة لشخص باستخدام صورتين مختلفتين.**




- **تم دمج صورة الشخصين (a) و (c) في الصورة الجديدة (b). المصدر: .(Huijstee et al.2021. p.10)**

ب- التلاعب بالوجه، مثل تعديل تعابيره ومزامنة الشفاه باستخدام الشبكات التوليدية التنافسية. تستخدم هذه الطريقة خوارزميتين للذكاء الاصطناعي، حيث يتم إدخال بيانات عشوائية في الخوارزمية الأولى، تعرف باسم خوارزمية التوليد لتحويلها إلى صورة، ثم تضاف هذه الصورة المصطنعة ضمن سلسلة من الصور الحقيقية لبعض المشاهير على سبيل المثال، ويتم إدخالها في الخوارزمية الثانية المعروفة باسم خوارزمية التمييز (Discriminator) وتوليد الصور الجديدة (Generator). وبعد تنفيذ عدد من الدورات والملاحظات، تبدأ الخوارزمية في إنتاج وجوه واقعية تماماً لأشخاص غير حقيقيين، كما يتضح من الشكل رقم: 3 أدناه.

**الشكل رقم3: يوضح عملية التلاعب بتعابير الوجه.**

- **يستخدم فيديو حقيقي (أعلى يسار الصورة) كمصدر للعملية، والصورة المستهدفة بالتعديل (صورة الرئيس الأمريكي السابق: ترامب) هي أسفل يسار الصورة العامة، والتي يتم تحويلها إلى فيديو حقيقي كما هو موضح على يمين الصورة العامة. (المصدر: (Huijstee et al. 2021. p10).**

**ثانياً- المحتوى الصوتي:** ويقصد به وبشكل رئيسي، تركيب الصوت وتعديله، إما عن طريق إنشاء ملف صوتي يتضمن حديثاً مزيفاً بنفس صوت الشخص، لكنه لم يقله في الحقيقة، أو عن طريق التحكم بنبرة صوت الشخص لإظهار شعور أو سلوك غير حقيقي.

**4-3- تهديدات التزييف العميق ومخاطره:**

وفقاً لتقرير صادر عن جامعة لندن العالمية (UCL)، يعد "التزييف العميق" لمقاطع الفيديو أو الصوت أكثر استخدامات الذكاء الاصطناعي إثارة للقلق ويستخدم في ارتكاب الجرائم والإرهاب (UCL. 2020. At: https://www.ucl.ac.uk/news/2020/aug/deepfakes-ranked-most-serious-ai-crime-threat). التقرير الآنف ذكره، استند على دراسة نُشرت في دورية "علم الجريمة" (Crime Science) وتم تمويلها من قبل مركز (Dawes) للجريمة المستقبلية في جامعة كاليفورنيا بالولايات المتحدة، حيث حددت الدراسة عشرون طريقة يمكن من خلالها استخدام الذكاء الاصطناعي لتسهيل الجريمة على مدار الـ15 عاماً القادمة، وتم تصنيفها من قبل 31 خبيراً في الذكاء الاصطناعي، بناء على الضرر الذي يمكن أن تسببه، وإمكانية الربح أو الكسب الإجرامي، ومدى سهولة تنفيذها، ومدى صعوبة إيقافها (Ibid).

وبناء على نتائج تلك الدراسة، فإن مقاطع الفيديو التي تم إنشاؤها بواسطة الذكاء الاصطناعي بالاعتماد على تقنية "التزييف العميق" (Deepfakes) لأشخاص حقيقيين يفعلون أشياء خيالية، قد حلت بالمركز الأول لسببين رئيسين (Ibid):

الأول- من الصعب تحديدها ومنعها. فلا تزال طرق الاكتشاف الآلي غير موثوقة، كما أن التزييف العميق يتحسن أيضاً في خداع عيون البشر. وقد أدت منافسة حديثة على موقع "فيسبوك" لاكتشافها باستخدام الخوارزميات إلى اعتراف الباحثين بأنها "مشكلة لم يتم حلها إلى حد كبير".

الثاني- يمكن استخدام تقنية "التزييف العميق" في مجموعة من الجرائم والأفعال السيئة، من تشويه سمعة الشخصيات العامة إلى عمليات النصب على الناس من خلال انتحال شخصيات أناس لهم علاقة بتلك الأموال.

وفي تقرير ذو صلة بهذا الموضوع، صدر في شهر يوليو-جويليه 2021 أعدته خدمة البحث في البرلمان الأوروبي (EPRS)، مستنداً على دراسة تناولت الجوانب التقنية والمجتمعية والتنظيمية لتقنية "التزييف العميق"، حددت هذه الدراسة عدداً من الأضرار التي يمكن أن يحدثها "التزييف العميق"، حيث تم تصنيفها في ثلاثة أقسام، هي: أضرار نفسية، أضرار مالية، وأضرار اجتماعية، تفاصيلها على النحو التالي: (Huijstee et al. 2021. P.iv:)

**الجدول رقم1: الأضرار التي يسببها التزييف العميق**

| الأضرار النفسية | الأضرار المالية | الأضرار الاجتماعية |
|---|---|---|
| ▪ الابتزاز(الشخصي)<br>▪ القذف<br>▪ التخويف<br>▪ التنمّر<br>▪ تقويض الثقة | ▪ الابتزاز(المالي)<br>▪ سرقة الهوية<br>▪ الاحتيال (مثل التأمين/الدفع)<br>▪ التلاعب بأسعار الأسهم<br>▪ تدمير العلامة التجارية<br>▪ الإضرار بالسمعة | ▪ التلاعب بالمحتوى الإعلامي<br>▪ زعزعة الاستقرار الاقتصادي<br>▪ تدمير لنظام العدالة<br>▪ تدمير للنظام العلمي<br>▪ انخفاض مستوى الثقة<br>▪ تدمير للديمقراطية<br>▪ التلاعب بالانتخابات<br>▪ زعزعة العلاقات الدولية<br>▪ تدمير للأمن الوطني |

- **المصدر: (Huijstee et al. 2021.p.iv).**

**4-4- الحلول والاقتراحات لمواجهة التزييف العميق للمحتوى:**

من الممكن اكتشاف التزييف العميق (Deepfakes)، إلا أن هناك بعض أنظمة الذكاء الاصطناعي المتطورة التي يمكن استخدامها لجعل المحتوى المزيف أقرب إلى الواقع، مما يزيد من صعوبة اكتشافه. ويمكن اكتشاف معظم مقاطع الفيديو المزيَفة من خلال ما يلي (مجلس جودة الحياة الرقمية. مرجع سابق. ص 13):

1. حركات الشخص الفوضوية وغير المنتظمة.
2. حدوث تغيّر مفاجئ في الإضاءة الموجهة للشخص.
3. تغيّر لون البشرة أثناء المقطع.
4. رمش العين بشكل متكرر أو عدم رمشها على الإطلاق.
5. عدم تطابق حركة الشفاه مع الكلام المسموع.
6. تشوه المنطقة المحيطة بالوجه.

ويقدّم خبراء في "مركز الأمن والتكنولوجيا الناشئة" ( CENTER for SECURETY and EMERGING TECHNOLOGY) والمعروف اختصاراً بالـ (CSET) التابع لجامعة "جورج تاون" بالولايات المتحدة، في أحد

تقاريره والذي جاء تحت عنوان "التزييف العميق: تقييم التهديد الأساسي" عدد من التوصيات المتعلقة بمواجهة "التزييف العميق"، أهمها، ما يلي (Hwang. 2020. PP.24-26):

1. بناء "حديقة" التزييف العميق (Deepfakes Zoo): يتم من خلالها تحديد عمليات "التزييف العميق" المرتبطة بالوصول السريع إلى نماذج من المحتوى الإعلامي "المزيف" والتي يمكن استخدامها لتحسين خوارزميات الكشف. فيجب على المنصات والباحثين والشركات الاستثمار في إنشاء "حديقة التزييف العميق" يقومون من خلالها بتجميع قاعدة بيانات خاصة بالمحتوى الإعلامي "المزيف" وإتاحتها مجاناً عند ظهورها على الانترنت.
2. تشجيع تتبع أفضل للقدرات (Encourage Better Capabilities Tracking): إذ توفر الأدبيات التقنية حول موضوع "تعلّم الآلة" (ML) نظرة ثاقبة حول كيفية استخدام الجهات الفاعلة في مجال المعلومات المضللة للتزييف العميق في عملياتها والقيود التي قد تواجهها في القيام بذلك. ومع ذلك، فإن ممارسات التوثيق غير المتسقة بين الباحثين تعيق هذا التحليل. يجب أن تعمل مجتمعات البحث ومؤسسات التمويل والناشرون والأكاديميون على تطوير معايير مشتركة للإبلاغ عن التقدم المحرز في النماذج التوليدية [iv].
3. كشف عملية "التسليع"(Commodify Detection): يمكن أن يؤدي توزيع تكنولوجيا الكشف على نطاق واسع إلى إعاقة فعالية تقنية التزييف العميق. يجب على الوكالات الحكومية والمنظمات الخيرية توزيع المنح للمساعدة في ترجمة نتائج الأبحاث في اكتشاف التزييف العميق إلى تطبيقات سهلة الاستخدام لتحليل مضمون وسائل الإعلام.

كما أن الدورات التدريبية المنتظمة للصحفيين ولأصحاب المهن التي من المحتمل أن تستهدفها هذه الأنواع من التقنيات قد تحد أيضاً من مدى إمكانية خداع أفراد من الجمهور.

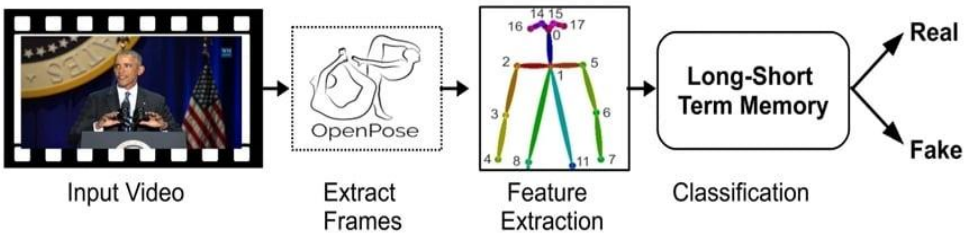
4. نشر "البيانات المشعة"(Proliferate Radioactive Data): أظهرت الأبحاث الحديثة أنه يمكن جعل مجموعة البيانات (Datasets) مشعة. حيث تقوم أنظمة "تعلم الآلة" (ML) المدربة على هذا النوع من البيانات بتوليد محتوى إعلامي مزيف، يمكن التعرف عليه بسهولة.

قد تبدو هذه التوصيات صعبة الاستيعاب على القارئ العادي أو غير المتخصص، لكنها توصيات عالية المستوى، تقدم حلولاً نوعية، موجهة بشكل خاص إلى المتخصصين في حقل المعلوماتية بالدرجة الأولى، ثم إلى الفاعلين الآخرين في المجتمع (المجتمع الغربي بالدرجة الأولى) الذين لهم صلات بهؤلاء المتخصصين، والذين يسعون جميعاً لإيجاد حلول لتلك المشكلة التي باتت تؤرق الحكومات في المقام الأول.

كما يقدم ثلاثة من الباحثين في مقال علمي مشترك لهم تحت عنوان:" محاربة التزييف العميق باستخدام تحليل لغة الجسد" طريقة للكشف عن التزييف العميق باستخدام تحليل لغة الجزء العلوي من الجسم، على وجه التحديد. حيث قاموا بتصميم شبكة (LSTM) متعددة الأطراف، وتم تدريبها كأنموذج تصنيف لاكتشاف التزييف العميق. بعد ذلك، قاموا بتدريب نماذج مختلفة من خلال تغيير "المعلمات الفائقة"

(hyperparameters) لبناء أنموذج نهائي بدقة معيارية، وقد حققوا درجة دقة تقدر بـ 94.39% في مجموعة اختبارات كشف التزييف العميق. وأظهرت النتائج التجريبية التي قاموا بها، أن لغة الجزء العلوي من الجسم يمكن أن تكشف بشكل فعال عن التزييف العميق (Yasrab et al. 2021. PP.303-321).

**الشكل رقم 4: يوضح أنموذج (LSTM) لكشف "التزييف العميق".**




- **في هذا الأنموذج، تتم عملية اكتشاف التزييف العميق من خلال استخراج إطارات الفيديو، ثم يتم تدريب خوارزمية التعلم الآلي عليها، وفي مرحلة لاحقة، يقوم نظام الاستدلال بتصنيف الفيديو إن كان مزيفاً أو حقيقياً.**

**المصدر: (Yasrab et al. 2021. PP.304).**

## خاتمة وعرض للنتائج:

من خلال ما تقدم، تتضح لنا التهديدات التي قد يمثلها التوظيف السلبي لبعض تقنيات الذكاء الاصطناعي، حيث دق العلماء والباحثين والمتخصصين ناقوس الخطر، إلى درجة أن بعض المتشائمين من أهل الاختصاص على غرار "إيلون ماسك" رئيس شركة "سبيس إكس" الأمريكية لتصنيع مركبات الفضاء، وعالم الفيزياء البريطاني الشهير "ستيفن هوكينغ" يريان أن الذكاء الاصطناعي يشكل تهديداً للبشرية.

لا يجب علينا التهوين من مخاطر وتهديدات تقنيات الذكاء الاصطناعي، كما لا يجب علينا التأخر في الاستفادة من تقنياته المختلفة التي يمكن توظيفها بشكل إيجابي ومفيد في مختلف شؤون حياتنا كما هو الحال مع استخدامنا لها في أجهزتنا الذكية المختلفة كالهواتف، والساعات، وأجهزة التلفاز، وغيرها.

أما التزييف العميق للوسائط الإعلامية، فبالإضافة للتوصيات والحلول التي قدمها الخبراء في مواجهة ذلك الخطر، فإنه يمكن أيضاً أن نقدم اقتراحات تسهم في الحد من خطر تلك التقنية على مجتمعاتنا العربية، لعل من أهمها، تدريب شباب متخصصين في مجال الذكاء الاصطناعي بشكل عام، والتزييف العميق بشكل خاص. كما يمكن جلب برامج خاصة بكشف المحتوى الإعلامي المزيف من الدول الصديقة المتقدمة في هذا المجال.

وأما عن أهم النتائج التي توصلنا لها من خلال هذا البحث، فهي التالية:

1. يوجد تهديد حقيقي للاستخدام السلبي لتقنيات الذكاء الاصطناعي على مستقبل البشر، خاصة فيما يتعلّق بالجانب العسكري للذكاء الاصطناعي، وعلى وجه الخصوص منه، أنظمة الأسلحة ذاتية التشغيل.

2. يحذر الباحثون والمتخصصون في علم الذكاء الاصطناعي من التهديدات والمخاطر المختلفة التي قد تسببها تقنية "التزييف العميق" على الأفراد والمجتمعات والحكومات بشكل عام، وفي عدة مجالات.
3. يمكن الحد من مخاطر التزييف العميق، من خلال برامج تعتمد على الذكاء الاصطناعي، إلا أن المشكلة التي قد يتم مواجهتها في هذا الشأن، هو أن مستخدمي تكنولوجيا "التزييف العميق" يطورون أساليبهم التقنية باستمرار.
4. يمكن التقليل من تهديدات "التزييف العميق" أيضاً، من خلال التقليل من نشر الصور الشخصية على شبكة الإنترنت عموماً، وعلى وسائل التواصل الاجتماعي بشكل خاص.

وأخيراً، فإن ما يجدر الإشارة إليه هنا أيضاً، هو أن تقنية "التزييف العميق" يمكن توظيفها بشكل إيجابي في عدد من المجالات، منها المجال الطبي، وفي مجال الترفيه، وفي مجال خدمة العملاء، وغيرها من الاستخدامات المفيدة، لكن لا يتسع المجال هنا لذكر تفاصيلها، ولكن يمكن تناولها في مقال آخر في المستقبل القريب.

- **قائمة المراجع**:


1- أوسوبا أوسوندي أ.، وويسلر الرابع ويليام. (2017). **مخاطر الذكاء الاصطناعي على الأمن ومستقبل العمل**. (ترجمة مؤسسة Rand). الولايات المتحدة: مؤسسة Rand.

2- عبد الوهاب شادي وآخرون. (2018). **فرص وتهديدات الذكاء الاصطناعي في السنوات العشر القادمة**. الإمارات: مركز المستقبل للأبحاث والدراسات المتقدمة، دورية اتجاهات الأحداث، العدد 27، (تقرير المستقبل: ملحق خاص يصدر مع دورية اتجاهات الأحداث).

3- قاسم علي. (2020). **التزييف العميق.. الجانب المظلم للذكاء الاصطناعي**. جريدة العرب (الدولية). العدد 11763.

4- مجدي نرمين. (2020). **الذكاء الاصطناعي وتعلم الآلة**. الإمارات: صندوق النقد العربي.

5- مجلس جودة الحياة الرقمية. (2021). **دليل التزييف العميق**. الإمارات: مجلس جودة الحياة الرقمية.

6- موسى عبد الله، وبلال أحمد حبيب. (2019). **الذكاء الاصطناعي: ثورة في تقنيات العصر**. القاهرة: المجموعة العربية للتدريب والنشر.

7- ندى بدر جراح. (2019). **تقنيات الذكاء الاصطناعي لتطوير التعلم الآلي الإحصائي**. العراق: المجلة العراقية لتكنلوجيا المعلومات، المجلد 9، العدد 3.

8- ويتباي بلاي. (2008). **الذكاء الاصطناعي: دليل المبتدئين**. (ترجمة: قسم الترجمة بدار الفاروق). القاهرة: دار الفاروق للاستثمارات الثقافية.

9- Huijstee Mariette van et al. (2021). **Tackling deepfakes in European Policy**. European Union : EPRS.

10- Hwang Tim. (2020). **Deepfakes: A Grounded Threat Assessment** (Report). USA: Georgetown's University.

11- Persons Timothy M. (2020). **Deepfakes**. USA: GAO.

12- Westerlund Mika. (2019). **The Emergence of Deepfake Technology: A review**. Canada: Carleton University, Technology Innovation Management Review, Vol 9, (Issue 11).

13- Yasrab Robail, Jiang Wanqi, Riaz Adnan. (2021). **Fighting Deepfakes Using Body Language Analysis**. Switzerland: MDPL.

14- [https://www.ucl.ac.uk/news/2020/aug/deepfakes-ranked-most-serious-ai-crime-threat].


(i)- أستاذ محاضر حول العلم المعرفي والذكاء الاصطناعي في إحدى الجامعات في بريطانيا. درّس تكنولوجيا الذكاء الاصطناعي لما يقرب من عقدين من الزمن، وتركزت أبحاثه حول النتائج الاجتماعية والأخلاقية لهذه التكنولوجيا الجديدة. كما يعمل "ويتباي" رئيساً لتحرير الجريدة الرئيسة المعنية بالذكاء الاصطناعي في بريطانيا.

(ii)- الأتمتة: هي عبارة عن استخدام الآلة في الأنشطة العادية، بحيث تصبح آلية أو تعمل بشكل تلقائي اعتماداً على الآلات. (الباحث).

(iii)- التعلم الآلي: هو أحد أنواع الذكاء الاصطناعي (AI) الذي يسمح للتطبيقات البرمجية أن تصبح أكثر دقة في تنبؤ النتائج دون القيام ببرمجتها بشكل صريح، ويمكننا بناء آلات لمعالجة البيانات والتعلم من تلقاء أنفسنا دون الإشراف المستمر. (يرجع في ذلك إلى: ندى بدر جراح، تقنيات الذكاء الاصطناعي لتطوير التعلم الآلي الإحصائي. العراق: المجلة العراقية لتكنلوجيا المعلومات، المجلد 9، العدد 3، 2019، ص 45).

(iv)- النماذج التوليدية (Generative Models): هي استخدام تقنيات الذكاء الاصطناعي والإحصاء والاحتمالات في التطبيقات لتوليد أو إنتاج أمثلة جديدة مستوحاة من الأمثلة الموجودة في مجموعة البيانات الأصلية. تستخدم النمذجة التوليدية في التعلم الآلي، وتحديداً التعلم غير الموجه كوسيلة لوصف الظواهر في البيانات بهدف تمكين الحواسيب من فهم العالم الحقيقي. (لمزيد من التفاصيل حول هذا المفهوم، يرجع إلى: (الهيئة العليا للعلوم والتكنولوجيا والابتكار- اليمن- على الرابط التالي: http://hasti.gov.ye[]، بتاريخ: 3-2-2022 على: 10:15م.