Journal of Police and Legal Sciences

Volume 15 | Issue 2 Article 2

7-8-2024

The Dangers of Terrorist Organizations Using Deepfakes and Ways to Confront

Mohammad Bedeir

Follow this and additional works at: https://www.jpsa.ac.ae/journal

 Part of the Critical and Cultural Studies Commons, Gender, Race, Sexuality, and Ethnicity in Communication Commons, International Humanitarian Law Commons, Law and Politics Commons, Law and Society Commons, and the Social Influence and Political Communication Commons

Recommended Citation

Bedeir, Mohammad (2024) "The Dangers of Terrorist Organizations Using Deepfakes and Ways to Confront," *Journal of Police and Legal Sciences*: Vol. 15: Iss. 2, Article 2.
DOI: https://doi.org/10.69672/3007-3529.1028

This Article is brought to you for free and open access by Journal of Police and Legal Sciences. It has been accepted for inclusion in Journal of Police and Legal Sciences by an authorized editor of Journal of Police and Legal Sciences. For more information, please contact Uq2012@hotmail.com.

مخاطر استخدام التنظيمات الإرهابية للتزييف العميق وسبل المواجهة

د. محمد بدرت بدير

الملخص:

زاد الاعتماد المجتمعي على البيانات الرقمية والتكنولوجيا السيبرانية، وحدثت طفرة مذهلة في مجال تقنيات الذكاء الاصطناعي في السنوات الأخيرة، كان من نتاجها اندماج سريع للذكاء الاصطناعي في الحياة اليومية، من خلال الأجهزة الذكية والمدن الذكية، وعلى الرغم من أن هذا ينطوي على العديد من المزايا، حيث يتم استخدامها في صناعات الترفيه والإعلام، إلا أنه يمثل أيضاً ضعفاً متزايداً أمام الهجمات السيبرانية، والتي قد تتجلى في هجمات التزييف العميق.

واعتمدَ الباحث على المنهج الوصفي التحليلي، وذلك بهدف وصف الظاهرة قيد البحث، ودراستها من مختلف أبعادها وجوانبها، ومن ثم تحليل مفرداتها ومكوناتها تحليلاً علمياً وعملياً، من أجل بلورة رؤية جديدة شاملة ومرنة للوقاية من هجمات التزييف العميق والحد من تداعياتها السلبية.

وأكد الباحث في نتائجه أن هجمات التزييف العميق لن تقتصر آثارها التخريبية على مستوى الأفراد والجماعات، وإنما تمتد لتزعزع استقرار الدول، وتشيع الفوضى، وتثير البلبلة بين مختلف الشعوب والحكومات، فالعالم اليوم أصبح قرية كونية صغيرة بفضل سيطرة التكنولوجيا الرقمية.

وارتكز البحث في أهميته على إبراز أهم المخاطر المحتملة لاستخدام التنظيمات الإرهابية للتزييف العميق، فضلاً عن إرسائه لمنهجية علمية متطورة تقوم على الوقاية من هجمات التزييف العميق الحالية والمستقبلية.

الكلمات المفتاحية:

التزييف العميق – شبكة الخصومة التوليدية – خوارزميات التعلم العميق – الطب الشرعي الرقمي – العلامة المائية الرقمية.

The Dangers of Terrorist Organizations Using Deepfakes and Ways to Confront

Dr. Mohammed Badart Bdeir

Abstract

Societal reliance on digital data and cyber technology has increased, and an amazing boom in the field of artificial intelligence technologies has occurred in recent years, resulting in the rapid integration of artificial intelligence into daily life, through smart devices and smart cities, and although this has many advantages, as it is used in the entertainment and media industries, it also represents an increasing vulnerability to cyberattacks, which may manifest itself in deepfake attacks.

The researcher relied on the descriptive analytical approach, with the aim of describing the phenomenon under research, studying it from its various dimensions and aspects, and then analyzing its vocabulary and components scientifically and practically, in order to crystallize a new comprehensive and flexible vision to prevent deepfake attacks and reduce their negative repercussions.

The researcher stressed in his results that deepfake attacks will not be limited to their destructive effects at the level of individuals and groups, but extend to destabilize states, spread chaos and cause confusion among different peoples and governments, as the world today has become a small global village thanks to the control of digital technology.

The research was based on highlighting the most important potential risks of terrorist organizations using deepfakes, as well as establishing an advanced scientific methodology based on the prevention of current and future deepfakes attacks.

Keywords :

Deepfakes – Generative Adversarial Network – Deep Learning Algorithms – Digital Forensics – Digital Watermark.

https://www.jpsa.ac.ae/journal/vol15/iss2/2
DOI: 10.69672/3007-3529.1028

المقدمة:

إن التلاعب بمحتوى الوسائط المتعددة ليس ظاهرة مُستحدَثة، بيد أنه في الآونة الأخيرة، تطورت بشكل كبير، وحظيت باهتمام واسع النطاق، نظراً للتطور المذهل في تزييف الصور ومقاطع الفيديو، وذلك بفضل استخدام التقنيات المُبتَكَرَة للتَعَلُّم الآلي والذكاء الاصطناعي في توليد محتويات مرئية وصوتية مُزيَّفة على قدر عالٍ من الدقة والإحكام، حتى أصبحت جودة مقاطع الفيديو نابضة بالحياة بشكل متزايد، ويصعب اكتشاف المكونات الاصطناعية بالعين المجردة، وبات من الصعب الوصول إلى الحقيقة، وانعدمت الثقة في المعلومات، فأصبحنا نعيش عصر "ما بعد الحقيقة".

وقد استخدمت التنظيمات الإرهابية هذه التقنية بشكل ضار لتكثيف الهجمات الإرهابية، أو لتضخيم إمكاناتها المادية عن طريق تداولها للعملات الرقمية المشفرة، وكذا لتعزيز إمكاناتها البشرية، وذلك بنشر الدعاية المتطرفة، وتجنيد عناصر جدد، والتحريض على العنف، مستغلة في ذلك سرعة بث المعلومات المُضلِّلة عبر الفضاء الإلكتروني، وتصميمها بثمن بخس من خلال البرامج المتاحة مجاناً، والقدرة على استئجار قوة المعالجة من خلال الحوسبة السحابية.

وقد ظهر – في يوليو ٢٠٢٠م - فيديو مفبرك للرئيس الأميركي الراحل ريتشارد نيكسون وهو يعلن فشل مهمة "أبولو ١١" الأميركية في صعودها إلى سطح القمر، ووفاة رواد الفضاء الذين شاركوا فيها على سطح القمر، وفي ٢٠١٨م، انتشر فيديو مزيف للرئيس الأمريكي الأسبق بارك أوباما، على الإنترنت بشكل واسع، وكان يوجه انتقادات لاذعة للرئيس السابق "دونالد ترامب"، ويصفه بألفاظ نابية، ويعد إنتاج مثل هذه الفيديوهات وغيرها للشخصيات العامة، وهم يقولون أشياء بأصواتهم وهيئتهم، بشكل يبدو واقعياً للغاية، تحديًا كبيرًا للأفراد والمؤسسات، يُقصد من وراءه إحداث البلبلة، وإثارة الشكوك، ونشر الفوضى، وهذا ما تهدف إليه الجهات الضارة كالتنظيمات الإرهابية وغيرها.

وقد ذكرت مؤخراً مجلة "**تايمز أوف إسرائيل**"، تصريحات مايكل ماتياس - الرئيس التنفيذي لشركة الوضوح – وهي شركة إسرائيلية لمواجهة تحدي التزييف العميق، وأدعى أن حركة حماس وأنصارها استخدموا التزييف العميق بالفعل في حربهم الجارية ضد إسرائيل، باخْتِلاقهم لفيديوهات مزيفة عن قتلى وجرحى من الأطفال والنساء؛ لمحاولة التأثير على الرأي العام، وجذب استعطافه، والتلاعب بالأوضاع السياسية.

إشكاليّة البحث:

تتمثل إشكاليّة البحث في الخطورة التي باتت تشكلها ظاهرة التزييف العميق، وقدرة التنظيمات الإرهابية على تذليلها لخدمة أغراضها المتطرفة والمنحرفة، وبث سمومها الخبيثة من أجل نشر دعوتها الغلواء، وتجنيد عناصر جدد، فضلاً عن استخدامها كسلاح من أسلحة الحرب النفسية؛ إذ تحرض - من خلال ترويج الفيديوهات المزيفة - على العنف، ونشر الفوضى، وتثبيط العزائم، وإحباط المعنويات، وتفكيك النسيج المجتمعي، وتمزيق الوحدة الوطنية.

وازداد الأمر خطورة، بسبب سهولة وسرعة انتشار محتويات الوسائط المتعددة المزيّفة عبر التطبيقات المجانية، ما أحدث تأثيراً أكثر تدميراً على المجتمعات بمختلف شرائحها وأطيافها، حيث باتت تؤثر في عقول الناس وتصرفاتهم، الأمر أدى إلى خلق تحديات أمام العالم أجمع، وهدّد السلِمَ المجتمعي، فهي ظاهرة يصعب كبح جماحها في ظل التطور التكنولوجي الهائل والمتنامي.

أهمية البحث:

أ. أهمية البحث النظرية:

تتجلى الأهمية النظرية للبحث ببيان الباحث للمفهوم العلمي لمصطلح التزييف العميق، وإبراز أنواعه المتعددة، كما تتمثل في بيان أهم الأساليب العلمية المُستحدَثة لتوليد هجمات التزييف العميق، والتي تَرَكَّزت في التشفير التلقائي، وشبكات الخصومة التوليدية (GANs).

ب. أهمية البحث العملية:

تبرز الأهمية العملية للبحث في تسليط الباحث الضوء على أهم المخاطر والتهديدات المحتملة لاستخدام التنظيمات المتطرفة للتزييف العميق، والتي قد تتمثل في التهديدات السيبرانية، والاستخباراتية، والمجتمعية، والأمن القومي، بغرض تحليل الظاهرة في سياقها العلمي والمجتمعي، والوقوف على مسار تطورها، من أجل التصدي لها، ومحاولة السيطرة على تعاظم آثارها.

وقد طرَح الباحث نهجاً علمياً وقائياً متعدد المحاور؛ لرصد ومنع هجمات التزييف العميق، واختتم دراسته ببعض التوصيات، التي يُرى أنها فاعلة ومؤثرة في الوقاية من هجمات التزييف العميق الحالية والمستقبلية.

أهداف البحث:

يهدف الباحث من خلال طرحه هذا الموضوع إلى تحقيق مجموعة مترابطة من الأهداف، تتمثل في الآتي:

1. التعريف بمفهوم التزييف العميق، وبيان تصنيفاته المتنوعة، والتي تنقسم إلى: نصي، صوتي، مرئي، مقاطع فيديو.
2. إيضاح الطرائق والوسائل التوليدية لخلق وإنشاء محتويات مزيفة من الوسائط المتعددة، وبيان مدى استطاعة خوارزميات التعلم الآلي المُستحدَثة والمتطورة على إنتاج وسائط متعددة عالية الدقة؛ إذ تحاكي الواقع.
3. إبانة كيفية استغلال التنظيمات المتطرفة للظاهرة محل البحث، وجعلها إحدى وسائلها في الحرب الإلكترونية، حيث تستهدف من خلالها تدمير البنى التحتية الحيوية، ونشر الأفكار المتطرفة، فضلاً عن إثارة البلبلة والشكوك، وتحطيم المعنويات والهمم، من أجل إحداث الفُرقة المجتمعية، وغير ذلك من الأغراض الإرهابية.
4. تسليط الضوء على أبرز التقنيات الحديثة لرصد ومنع هجمات التزييف العميق، وبيان أهم البروتوكولات الخاصة بالأمن السيبراني لمحاولة الوقاية من مخاطرها.

تساؤلات البحث:

يحاول الباحث من خلال طرحه لهذا البحث الإجابة على عدة تساؤلات هامة، نذكر منها ما يلي:

1. كيف تستطيع التنظيمات الإرهابية - عن طريق تطويعها لتقنية التزييف العميق - خرق الأنظمة الأكثر أماناً في الدول، والاضطلاع بأعمال التجسس؟
2. هل يمكن للجهات الضارة توظيف تقنيات التزييف العميق في تمزيق النسيج المجتمعي، وبث الفرقة وروح الكراهية بين صفوف المجتمع؟
3. هل يمكن الإضرار بالأوضاع والمصالح الاقتصادية للدول، بواسطة إنشاء مقاطع فيديو مزيفة وصور مُضلِّلة؟

https://www.jpsa.ac.ae/journal/vol15/iss2/2
DOI: 10.69672/3007-3529.1028

4. كيف يمكن للتنظيمات الإرهابية نشر أيدولوجيتها المتطرفة، وتجنيد عناصر جدد، باستخدام التزييف العميق؟
5. كيف تستغل الجماعات المتطرفة تقنيات التزييف العميق، في إضعاف الثقة في مؤسسات الدولة ورموزها؟
6. ما هي أبرز بروتوكولات الأمن السيبراني للرصد والاستجابة؟
7. كيف يمكن تحصين الدفاعات السلوكية، ضد مخاطر التزييف العميق؟
8. هل هناك أساليب تقنية مستحدثة يمكن استخدامها من أجل المنع والوقاية من التزييف العميق، وليس مجرد الكشف؟

فروض البحث:

تقوم فرضية البحث على إثبات فكرة أساسية تتمثل في أنه كلما تسابقت الأبحاث العلمية في تطوير تقنيات الذكاء الاصطناعي ونُظُم التَعَلُّم العميق، وهيمنت التكنولوجيا الرقمية على الحدود والمسافات والمستويات الفكرية والاقتصادية، وبسطت مواقع التواصل الاجتماعي سلطانها، زاد مستوى التحدي والتهديد لمخاطر هجمات التزييف العميق على أبعاد الأمن القومي بمختلف مستوياته، وانعكس ذلك بالضرورة على أساليب مواجهة هذا التهديد، الذي بات أشد فتكاً من الحروب التقليدية.

منهج البحث:

اعتمد الباحث بصورة رئيسة على المنهج الوصفي التحليلي، حيث يعتبر هذا المنهج طريقة من طرق التحليل والتفسير بشكل علمي مُنظَّم، وذلك بهدف وصف الحقائق والمعلومات المرتبطة بموضوع البحث، والعمل على تحليل مفرداتها ومكوناتها تحليلاً علمياً وعملياً، ودراستها من مختلف أبعادها وجوانبها؛ لاستخلاص أهم القواعد والأحكام التي ترتبط بالموضوع، حيث سلط الباحث الضوء على المخاطر المحتملة لاستخدام التنظيمات الإرهابية تقنيات التزييف العميق، بغرض الخروج بنتائج جوهرية، ومن ثم بلورة رؤية جديدة شاملة ومرنة لدحض ومنع هجمات التزييف العميق.

وقد اعتمد الباحث في هذا المنهج على المراجع العلمية المختلفة من كتب ودراسات وأبحاث، سواء باللغة العربية أو اللغة الأجنبية التي لها صلة بموضوع البحث.

الدراسات السابقة:

اضطلع الباحث برصد بعض الدراسات والأوراق البحثية ذات الصلة بموضوع الدراسة، وقد تم تصنيف الدراسات السابقة إلى ما يلي:

دراسة لـ(Nina Schick) لعام ٢٠٢٠م، بعنوان (Deepfakes: The Coming Infocalypse)، وقد رَكَّزَت الدراسة على الاستخدام والتأثير المُحتمَل للتزييف العميق، باعتباره نوعاً من الوسائط الاصطناعية التي تم إنشاؤها بواسطة الأجهزة الرقمية، كالصور ومقاطع الفيديو، وله القدرة على إنشاء تمثيلات مصطنعة لأفراد غير موجودين وإظهار أفراد حقيقيين يفعلون أشياء لم يفعلوها، وعلى هذا النحو، تُشكِّل تقنية التزييف العميق تهديداً واضحاً للتلاعب بالحقائق.

وأكدت الدراسة على أن تقنية التزييف العميق لا تشكل تهديداً حقيقياً للديمقراطية فحسب، بل تهدد الأمن القومي بكافة أبعاده، وتطرقت إلى تداعياته السياسية الخطيرة، وكيف يتم استخدام التزييف العميق للترهيب والانتقام والاحتيال، ومدى عدم استعداد الحكومات وشركات التكنولوجيا حقاً لما هو قادم.

كما استعرض الكتاب صوراً متعددة للجوانب الإيجابية للتزييف العميق، استكشف من خلالها الإمكانات التقنية للاستخدام الإبداعي في مجالات عدة كالأفلام، والإعلانات، والرسم، وفي النهاية تم وضع التزييف العميق في سياق ثقافي وفلسفي أوسع، مع التركيز بشكل أساس على فكر ما بعد الإنسانية.

دراسة لـ(Bryan Lyon، Matt Tora) لعام ٢٠٢٣م، بعنوان (Exploring Deepfakes)، أوضح فيها المؤلفان ماهية التزييف العميق وتاريخه، وتناولا التطبيقات الحديثة للتزييف العميق، وأنواعه المختلفة، والخوارزميات المستخدمة في توليد وإنشاء التزييف العميق، واستعرضا التحديات والمشاكل الحديثة باعتبار أن التزوير والتعديل الرقمي لمقاطع الفيديو والصور وكذلك المحتويات النصية أصبحت منتشرة ومتعددة، خاصة عندما يتم اعتماد تقنيات التزييف العميق في العديد من المصادر.

وسلط المؤلفان الضوء على كيفية استخدام الميزات الرئيسة لهذه التكنولوجيا بشكل أخلاقي، حيث بَيَّنَ المؤلفان القيمة التي يمكن أن تجلبها تقنية التزييف العميق إلى مجموعة متنوعة من حالات الاستخدام التعليمية والفنية، من الصور الرمزية لألعاب الفيديو إلى صناعة الأفلام.

دراسة لـ(Michael Filimowicz) لعام ٢٠٢٢م، بعنوان (Deep Fakes, Algorithms and Society)، وتناول المؤلف من خلال دراسته (التزييف العميق: الخوارزميات والمجتمع) طرائق استخدام تقنيات الذكاء الاصطناعي لإنتاج مقاطع سمعية بصرية واقعية وهمية، لا يمكن تمييزها عن وسائط الفيديو التقليدية.

وعقد المؤلف مقارنة بين الماضي والحاضر بشأن التقاط الصورة الفوتوغرافية، والوسائط المرتبطة بها من أفلام وفيديو (والتي كانت تجسد الحقيقة)، حيث كانت تنتج هذه الصور بتتبع سببي لأشعة الضوء الفعلية في زمان ومكان معينين، والتي يتم تثبيتها بواسطة التفاعلات الكيميائية أو أجهزة الاستشعار الرقمية للصورة الناتجة، أما اليوم، فيمكن إنشاء وسائط سمعية بصرية واقعية من شبكات التعلم العميق التي تقطع أي اتصال بحدث فعلي.

وطرح المؤلف تساؤلات عدة ومنها، هل يجب على المجتمع إنشاء أنظمة جديدة لإدارة هذا التحدي الجديد لإحساسنا بالواقع والقدرات الإثباتية التقليدية لـ"الصورة الميكانيكية"؟ وكيف تُولّد هذه الصور اضطراب المعلومات بينما توفر أيضاً الأساس للأدوات المشروعة المستخدمة في صناعات الترفيه والإبداع؟

وأبرز ما يميز هذه الدراسة عن الدراسات السابقة هي محاولة تركيز الباحث على سبل استخدام التنظيمات الإرهابية لتقنيات الذكاء الاصطناعي، وقدرتها على تسخير هجمات التزييف العميق لخدمة أغراضها المتطرفة، والنيل من مؤسسات الدولة، وتشوية صور الرموز والقادة والزعماء؛ للإضرار بمصالح الأمن الوطني للدول المستهدفة، فضلاً عن طرح الباحث لبعض التدابير الأمنية والتقنية اللازمة لمواجهة تلك المخاطر والحد من انتشارها، بالإضافة إلى المواجهة التشريعية على الصعيدين الدولي والإقليمي.

خطة البحث:

تم تقسيم خطة البحث وفقاً للآتي:

المبحث الأول: ماهيّة التزييف العميق

المطلب الأول: مفهوم التزييف العميق

https://www.jpsa.ac.ae/journal/vol15/iss2/2
DOI: 10.69672/3007-3529.1028

المبحث الأول

ماهيّة التزييف العميق

تمهيد وتقسيم:

نستعرض في هذا المبحث المفهوم العلمي للتزييف العميق، ثم نلقي الضوء على أهم أنواعه المختلفة ومنها: النص، الصورة، الصوت، مقاطع الفيديو، ثم ندلف لنبرز أهم الأساليب المستخدمة لتوليد التزييف العميق، وذلك كما يلي:

المطلب الأول- مفهوم التزييف العميق:

يشير التزييف العميق (Deepfakes) إلى التلاعب بمحتوى الوسائط المتعددة، حيث يتم تغييرها رقمياً أو إنشاؤها صناعياً من العدم؛ لاستبدال مظهر شخص حقيقي بآخر مزيف، أو التَفَوَّهَ بكلمات وألفاظ لم ينطق بها على الإطلاق، وذلك باستخدام الشبكات العصبية العميقة.

ويُعَرَّفَ بأنه نتاج المُخَادَعَة بالصوت أو الوجه، وتغيير سماته وتعبيراته، مثل: العمر، والجنس، ولون البشرة، ولون الشعر، ولون العين، والنظارات، والماكياج، والشارب، واللحية، والنظرة، والفم المفتوح أو المغلق، والإصابة، وغير ذلك (Maurizio, 2019).

وعَرَّفَه آخرون بأنه وسائط اصطناعية يتم إنتاجها باستخدام نماذج التَعَلُّم العميق، حيث تستبدل الأخيرة الميزات الموجودة في صورة واحدة بميزات صورة أخرى؛ لتبدو محتويات الوسائط السمعية والمرئية المفبركة كأنها حقيقية (Rajdeep, 2022).

ويرى البعض أن التزييف العميق مصطلح مُشتَقّ من التَعَلُّم العميق، ويتم استخدام خوارزميات التَعَلُّم العميق التي تُعلِّم نفسها كيفية حل المشكلات، مع مجموعات البيانات الكبيرة؛ لمبادلة الوجوه في مقاطع الفيديو والصور والمحتويات الرقمية الأخرى؛ لجعل المحتوى المُزيَّف يبدو أصلياً (Ian, 2021).

ويُعرِّف الباحث التزييف العميق بأنه محاكاة رقمية للواقع، ينتج عنها محتويات إلكترونية مُزيَّفة كالصور الرقمية، ومقاطع الفيديو، والموسيقى، وتعتمد على خوارزميات التَعَلُّم العميق، وتُستخدم إما لمبادلة المحتوى الرقمي بآخر مُستهدَف، أو للتلاعب بمظهره وهيكله الخارجي ليظهر على غير حقيقته، أو لإنشاء وتوليد محتوى جديد من نَظَائِره المُخزَّنة، ويتم ذلك باستخدام الطرائق التوليدية العميقة، والتي بات كشفها شبه مستحيل.

وبتطبيق تقنية التزييف العميق على الوجه البشري، يتضح أن هناك العديد من أساليب المخادعة، تتمثل في الآتي:

1. **تبديل الوجه:** هو الأسلوب الأكثر شيوعاً، حيث يتم استبدال وجه شخص ما في الصورة أو الفيديو المُستهدَف، بوجه شخص آخر في الصورة أو الفيديو الأصلي.
2. **التلاعب بتعبيرات الوجه:** يتطلب ذلك تغيير بعض تعبيرات الوجه لشخص ما في الصورة أو الفيديو المُستهدَف، بتعبير وجه شخص آخر في الصورة أو الفيديو الأصلي، كالحزن، والفرح، والابتسامة، والغضب، والارتباك، وغير ذلك من المشاعر والعواطف.
3. **معالجة سمات الوجه:** يتم ذلك باكتشاف السمات الحالية للوجه، ومحاولة تغيير بعض الميزات المحددة، مثل: لون الشعر، والبشرة، والعينين، والعمر، والجنس.
4. **توليد الوجه:** يتم ذلك بإنشاء عينات وجه غير موجودة، وبالتالي يتم توليد وجه لشخص لم يكن موجوداً من قبل (Maurizio, 2019).

مما سبق يتضح أن التزييف العميق يفرض واقعاً مغايراً، نظراً لأن محتواه المُزيَّف يبدو حقيقياً، ويسهل من خلاله خداع المشاهدين والمستمعين؛ للاعتقاد بأن ما يرونه أو يسمعونه حقيقي، وهو ما يؤدي إلى عواقب وخيمة، تضر بسمعة الأشخاص، وقد تصل إلى حد النيل من اقتصاد الدول وسياساتها.

المطلب الثاني- أنواع التزييف العميق:

نتناول فيما يلي التصنيفات المتعددة للتزييف العميق، والتي تنقسم إلى أربع مجموعات رئيسة، وهي كالتالي:

أ. التزييف العميق لمقاطع الفيديو:

تعد مقاطع الفيديو المُزيَّفة أكثر أنواع التزييف العميق خطورة؛ نظراً لسهولة إنشائها عن طريق برامج خلق وتعديل الفيديو، باستخدام تقنيات الذكاء الاصطناعي، فضلاً عن تأثيرها في العقول والمشاعر؛ لقدرتها على تجسيد الواقع ومحاكاته، وتسلسلها القصصي والمنطقي للأحداث، وتوصيلها الرسالة أو الهدف بشكل أفضل من التنسيق النصي المكتوب، أو الصورة (Rajdeep, 2022).

https://www.jpsa.ac.ae/journal/vol15/iss2/2
DOI: 10.69672/3007-3529.1028

وتتمثل خطورتها أيضاً في سرعة بثها، وقدرتها على الانتشار والتداول، بغرض الابتزاز والتشهير بالأشخاص التاريخيين، والمشاهير، والقادة، وغيرهم (Ignas, 2022).

وعلى سبيل المثال، انتشر مقطع فيديو - في عام ٢٠١٩م - على موقع الفيسبوك، يُظهر النائبة الأمريكية نانسي بيلوسي وهي تبدو في حالة سكر أثناء حديثها في مؤتمر صحفي، وذلك بعد أن خضع الفيديو لعملية دَبلَجة عالية الاحتراف؛ لجعل مظهرها وصوتها مشوهًا ومبطئًا، كما لو كانت تحت تأثير الكحول، وحقق الفيديو نسب عالية من المشاهدة والمشاركة (Ian, 2021).

ب. التزييف العميق للصوت:

أصبحت الجهات الضارة بإمكانها التلاعب بالأصوات ومقاطع الكلمات، حيث بات يسهل على خوارزميات التَعَلُّم الآلي اِستنساخ أي صوت بشري، ومحاكاته لنفس طبقة الصوت، واللهجة، والنبرة، والنغمة، وذلك بمجرد تَوَفُّر مقطع صوتي للشخص المراد تقليد صوته، وكلما زاد عدد التسجيلات الصوتية للشخص المُستهدَف، أصبح الصوت المُزيَّف أكثر دقة.

وعلى سبيل المثال، تظاهر مخادع - في عام ٢٠١٩م - من خلال محادثة هاتفية، بأنه الرئيس التنفيذي لإحدى شركات الطاقة الألمانية، وطلب تحويل مبلغ وقدره ٢٤٠ ألف يورو لحساب إحدى الشركات الموردة.

وفي عام ٢٠٢٣م، أعلنت شركة مايكروسوفت لتقنيات الحاسوب عن اكتشاف خوارزمية يمكنها إعادة إنتاج صوت بشري بناء على نموذج صوتي يبلغ طوله ثلاث ثوان فقط، وتعمل هذه التقنية بلغات متعددة، أي يستطيع الشخص سماع صوته يتحدث بلغات أجنبية، وذلك من دون أي جهد، سوى أن يقوم بتحميل تسجيل صوتي للصوت والكلمات المُستهدفة، أي التي سيتم التحدث بها.

ج. التزييف العميق للصور:

يستطيع أي فرد اِلتِقاط صور لأي شخصية من المشاهير أو غيرها، وبثها عبر الإنترنت بعد إخضاعها للتطبيقات والبرامج المختلفة والمجانية التي تستبدل وجوه الأشخاص، أو تتلاعب بالتعبيرات والملامح الرئيسة لتلك الوجوه، وتُولِّد صوراً مغايرة للحقيقة، ما يتسبب في العديد من الأضرار للشخصية المُنتهَكة، كالإضرار بالسمعة وغير ذلك (Ian, 2021).

د. التزييف العميق للنص:

لم يكتفِ مجرمو الإنترنت بالتزييف العميق للفيديوهات، والصوت، والصور، فقد أصبح التزييف العميق النصي ممكناً، وذلك بفضل تدريب الشبكة العصبية المتكررة[1] (RNN)، وقدرتها على معالجة اللغة الطبيعية (NLP)، حيث يمكن للبرامج التفكير في البيانات وتكييفها، واستخدامها في الكتابة والتأليف تماماً مثل البشر، وكتابة المقالات الإخبارية والقصائد والمدونات وغير ذلك (Floridi, 2020).

[1] الشبكة العصبية المتكررة Recurrent Neural Network (RNN): هي نوع من الشبكات العصبية الاصطناعية التي تستخدم بيانات متسلسلة، وتتميز بذاكرتها لأنها تأخذ المعلومات من المدخلات السابقة للتأثير على المدخلات والمخرجات الحالية، وتستخدم لترجمة اللغة، ومعالجة اللغة الطبيعية (nlp)، والتعرف على الكلام، والتعليق على الصور، ويتم دمجها في التطبيقات الشائعة مثل: Siri، والبحث الصوتي، وترجمة Google.

وفي إطار التطور المذهل لتقنيات التزييف العميق، أصبح الأخير يحدث بشكل مباشر، أي في الوقت الفعلي، حيث يمكن للمستخدمين الظهور بمظهر شخص آخر أثناء البث المباشر، أو في مكالمة الفيديو، وذلك عبر بعض البرامج الحديثة كـ(DeepFaceLive)، وهو برنامج ذكاء اصطناعي مفتوح المصدر يمكنه تغيير صورة أي شخص إلى صورة شخص آخر، عبر مؤتمرات الفيديو وشبكات البث (Nicholas, 2022).

المطلب الثالث- أدوات توليد التزييف العميق:

هناك العديد من الأساليب المستخدمة لتوليد التزييف العميق، وإنشاء فيديوهات مبادلة الوجوه، والتلاعب بتعابير الوجه البشري، وتغيير مظهره أو هويته بطرائق احترافية، بيد أن هناك أسلوبين رئيسيين لتوليد تلك التقنية، وهما: التشفير التلقائي، وشبكات الخصومة التوليدية، وذلك على النحو التالي:

أولاً- التشفير التلقائي (Autoencoder):

يعد التشفير التلقائي نوعًا خاصًا من خوارزمية التَعَلُّم العميق، فإذا كان لدينا صورة لوجه شخص ما، ونريد استبدالها بصورة لوجه شخص آخر، فيتم تدريب برنامج التشفير التلقائي للتزييف العميق على كل من الصورتين الأصلية والمستهدفة، ويمر ذلك بخطوتين، وهما كالآتي:

الخطوة الأولى: تضطلع الشبكة العصبية في هذه الخطوة بالتعرف على أوجه التشابه بين الوجهين الأصلي والمُستهدَف، وتقللها إلى تمثيلات أصغر (قائمة على الميزات المشتركة)، أي يتم ضغط الصور، ويكون ذلك بتشفير ميزات الوجهين مثل: (شكل الأنف، لون البشرة، لون العين، وما إلى ذلك) في مجموعة صغيرة من القيم العددية، حيث يتم الترميز من خلال سلسلة من الطبقات التي تبدأ بالعديد من المتغيرات، وتصبح أصغر تدريجياً حتى تصل إلى طبقة تسمى "عنق الزجاجة"، وتحتوي الأخيرة على العدد المُستهدَف من المتغيرات.

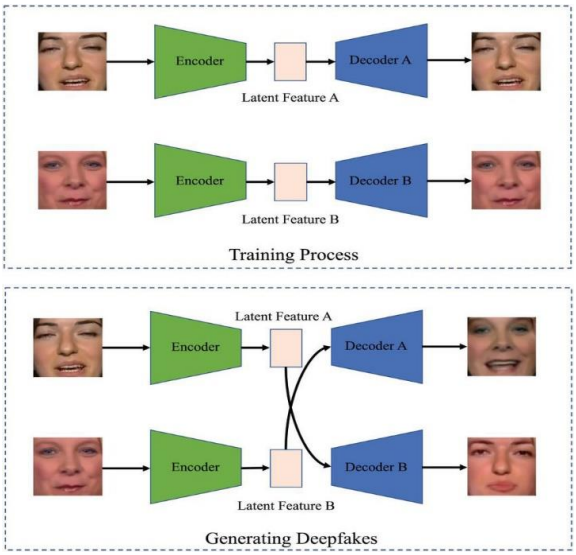
الخطوة الثانية: تقوم الشبكة العصبية بفك تشفير البيانات الموجودة في طبقة عنق الزجاجة؛ لاستعادة الوجوه من الصور المضغوطة، وإعادة إنشاء الصور الأصلية، حيث يتم تدريب وحدة فك الترميز الأولى لاستعادة وجه الشخص الأول، ووحدة فك الترميز الثانية لاستعادة وجه الشخص الثاني (Joan, 2022).

وأثناء عملية التدريب، يجب تزويد برنامج التشفير التلقائي بسلسلة من الصور، بغرض ضبط الأوزان في طبقات التشفير وفك التشفير، بحيث تكون صورة الإخراج مشابهة لصورة الإدخال قدر الإمكان، فكلما كانت مجموعة البيانات المُدخلة كبيرة، أصبحت نتائج برنامج التشفير التلقائي أكثر دقة (Xun, 2021).

وبعد تدريب أجهزة التشفير التلقائي، تقوم بتبديل مخرجاتها، أي يتم تبديل أجهزة فك التشفير، حيث يتم استخدام برنامج التشفير للصورة المصدر، ووحدة فك ترميز الصورة المُستهدَفة؛ لإخراج الأخيرة بميزات الأولى، وتصبح الصورة الناتجة لها وجه المصدر على وجه الهدف، مع الحفاظ على تعابير وجه الهدف، أو بمعنى آخر، يلتقط برنامج التشفير التلقائي تعبيرات وجه شخص ما، ويعينها على المظهر الخاص لوجه شخص آخر.

والشكل التالي يوضح كيفية إنشاء التزييف العميق، حيث يتم توصيل مجموعة ميزات الوجه (أ) بجهاز فك التشفير (ب)؛ لإعادة بناء الوجه (ب) من الوجه الأصلي (أ).

https://www.jpsa.ac.ae/journal/vol15/iss2/2
DOI: 10.69672/3007-3529.1028



ثانياً- شبكة الخصومة التوليدية (GAN):

Generative Adversarial Network

تعد شبكات الخصومة التوليدية (GAN) من أفضل الطرائق لتوليد صور واقعية للأشياء والمشاهد والأشخاص، فهي نموذج للتَعَلُّم الآلي غير الخاضع للإشراف، ولديها القدرة على اكتشاف وتعلم الأنماط في بيانات الإدخال تلقائياً، وذلك بتَّدَرُّبُها على إنشاء صور واقعية من الضوضاء.

وتُعرَّف شبكة (GAN) بأنها عبارة عن خوارزميتين للذكاء الاصطناعي يعملان ضد بعضهما البعض، أحدهما يسمى (المُوَلِّد)، وهو مُكَلَّف بإنشاء المحتوى الرقمي المُزيَّف، من خلال تغذيته بضوضاء عشوائية، والآخر يسمى (المُمَيِّز)، ويضطلع بالتمييز فيما بين المحتوى الحقيقي والمُصطنَع (Goodfellow, 2023).

البِنْيَة التكوينية لشبكة الخصومة التوليدية:

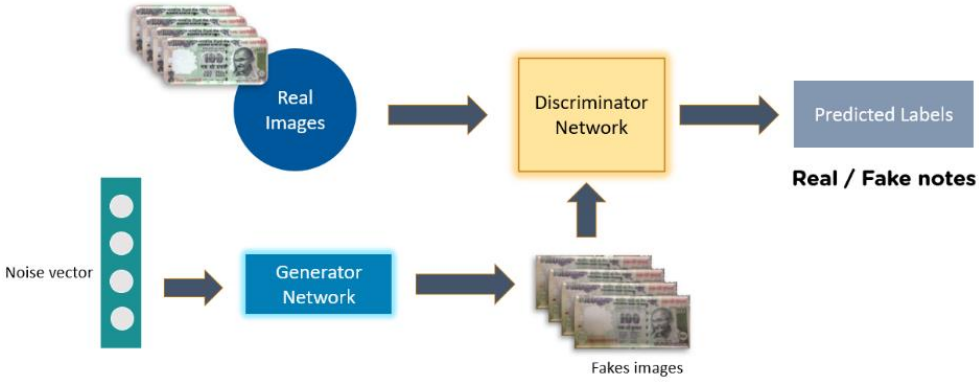
تتألف شبكة الخصومة التوليدية من مُكوَّنين رئيسين، وهما كالآتي:

أ. **المُوَلِّد:** عبارة عن شبكة عصبية غير خاضعة للرقابة، يتجلى دورها في إنشاء وتوليد بيانات أو عينات وهمية على أساس العينة الأصلية، حيث يتم تغذيتها بمدخلات عشوائية، قد تكون صورة أو نصًا أو صوتًا، ويتم تحسين أدائها على أساس التغذية المرتدة من المُمَيِّز، بهدف إنتاج عينات لا يستطيع الأخير تمييزها عن البيانات الحقيقية، وتعد العينات التي تم إنشاؤها أمثلة تدريب سلبية للمُمَيِّز (Mariëtte, 2023).

ب. **المُمَيِّز:** عبارة عن شبكة عصبية خاضعة للإشراف، يتم تغذيتها بالعينات الحقيقية، والعينات التي تم إنشاؤها بواسطة المُوَلِّد، ويتمثل دور المميز في تحديد التشوهات فيما بين العينات، بغرض تصنيف البيانات الحقيقية بشكل صحيح على أنها حقيقية، والبيانات التي تم توليدها على أنها مزيفة، ثم يوفر التغذية المرتدة للمُوَلِّد لتحسين أدائه.

ويستمر التدريب في شكل عدائي بين شبكتي "المُوَلِّد والمُمَيِّز"؛ لفحص الاختلافات والتقاطها وتكرارها داخل مجموعة البيانات، حيث تهدف الشبكة الأولى إلى توليد عينات خادعة للشبكة المُمَيِّزة، بينما تحاول الأخيرة تحسين قدرتها للتمييز بين البيانات الحقيقية والمفبركة (Ding, 2020).

وفي البداية، لن تبدو الصور الاصطناعية مثل الأصلية، ولكن مع تكرار هذه العملية، واستمرار التدريب، ومعرفة الناتج من خلال "**الانتشار العكسي**"[1]، والعمل على تحسين الأداء لكل من المُمَيِّز والمُوَلِّد، ينجح الأخير في إنتاج عينات واقعية تماماً كالحقيقية، بينما يصبح المُمَيِّز أكثر مهارة في الفرز والتصنيف بين أنواع البيانات الأصلية والمُزَيَّفة، إلى أن تتقارب هذه العملية إلى نقطة يكون فيها المُوَلِّد قادراً على توليد عينات عالية الجودة، يصعب على المُمَيِّز تمييزها عن البيانات الحقيقية، بل وإنتاج وجوه جديدة لأشخاص غير موجودة بالفعل.



المبحث الثاني

صورتهديدات التزييف العميق من جانب التنظيمات الإرهابية

تمهيد وتقسيم:

تمتد التداعيات والآثار الضارة للتزييف العميق إلى ما هو أبعد من الإضرار بالسمعة، حيث ظهرت - في الآونة الأخيرة - كأداة قوية في يد التنظيمات الإرهابية، تستطيع من خلالها استهداف البنى التحتية الحيوية، وتعطيل العمليات الحكومية، وإحداث الفوضى في الأسواق المالية، فضلاً عن نشر الأفكار المتطرفة، والتحريض على العنف، وغير ذلك.

وسيصبح الخطر الذي يجلبه التزييف العميق أكثر أهمية في المستقبل، نظراً للتَحَسُّن والتطور المستمر في أساليب التعلم الآلي المستخدمة بشكل أكبر، فضلاً عن سرعة بث المحتوى الرقمي للتزييف العميق عبر المواقع المجانية للتواصل الاجتماعي، وفيما يلي استعراض لأبرز مخاطر هجمات التزييف العميق، وهي كالآتي:

المطلب الأول- التهديدات السيبرانية:

تتزايد المخاوف إزاء استغلال التنظيمات الإرهابية لمواطن الضعف المتأصلة في الفضاء السيبراني والطبيعة غير المتكافئة للتهديدات التي تشكلها الهجمات السيبرانية، حيث تضطلع بأعمال التخريب عن طريق دمج تقنية التزييف العميق (Deepfakes) في حملات الهندسة الاجتماعية، أو بالتغلب على أنظمة الكابتشا[2] (CAPTCHA)؛ ليتم زرع برامج ضارة، أو محتوى إرهابي، أو مواد تخريبية، أو دعائية، وهو ما يعرض المعلومات للخطر، حيث تنتهك سريتها، وتهدد سلامتها وتوافرها (Nicholas,2022).

---

[1] الانتشار العكسي (Inverse propagation): هو الانتشار من طبقة الإخراج إلى الإدخال، فهو يعد تصحيحًا وتحديثًا لقيمة الأوزان، بناءاً على معدل الخطأ الذي أدى إلى التمرير الأمامي.

[2] اختبار CAPTCHA هو إجراء أمني مصمم لحماية الشبكات ومواقع الويب من الهجمات، حيث يضع أسئلة لا يستطيع حلها سوى عقل بشري قادر على التمييز، وبالتالي يمكن تمييز المستخدمين الحقيقيين عن الروبوتات، ومن ثم حظر الأخيرة.

https://www.jpsa.ac.ae/journal/vol15/iss2/2
DOI: 10.69672/3007-3529.1028

وتزداد قوة التهديدات السيبرانية حينما ترصد الجماعات الإرهابية يوم الصفر (Zero Days)، وهو مصطلح يستخدم لوصف خلل في بنية البرنامج، أو ثغرة أمنية في تطبيق أو أنظمة التشغيل، غير معروفة للمطورين، أو لا يوجد تصحيح أمان متاح لها، وبمجرد اكتشاف الجهات الضارة لتلك الاختلالات، وقبل تصويب مسارها، أو العمل على حلها (من قبل مقدمي الخدمة)، تضطلع هي باستهداف النظام، وذلك بحقنه بأنواع مختلفة من البرامج الضارة، ولعل من أبرزها ما يلي:

1. الخزانة العميقة (DeepLocker):

هي تقنية هجوم خفية شديدة الاستهداف والمراوغة، تضطلع من خلالها الجهات الفاعلة الضارة بإخفاء البرامج الخبيثة بشكل غير مرئي داخل تطبيق شرعي وشائع، وتبرز خطورته في أنه يصيب ملايين الأنظمة دون أن يتم اكتشافه على الإطلاق.

وحظر خبراء الأمن السيبراني - في أعقاب مؤتمر البلاك هات (Black Hat) بالولايات المتحدة الأمريكية لعام ٢٠١٨م - من استغلال الجماعات الإرهابية للتطور الهائل للبرنامج الضار (DeepLocker)، وقدرته على التنكر في تطبيق مؤتمرات فيديو، حيث يظل كامناً، حتى يحدد ضحيته المقصودة، من خلال التعرف على الوجه، أو الصوت، أو تحديد الموقع الجغرافي، ثم ينشر حمولته الضارة (Menno, 2022).

2. برامج الفدية:

هي برامج ضارة تقوم بتشفير ملفات الضحايا، بعد أن يتم الوصول إليها عن طريق التحايل بهجمات التزييف العميق، من أجل المطالبة بدفع مبالغ مالية كفدية مقابل فك تشفير الملفات، وهذا ما سعت إليه التنظيمات الإرهابية لتضخيم مصادر دخلها، حتى تتمكن من تنفيذ هجماتها الإرهابية سواء الرقمية أو المادية، وقد تجلى ذلك في هجوم (WannaCry) لعام ٢٠١٧م، حيث انتشر خلل في تطبيق (Microsoft Windows)، واستطاع تشفير جميع البيانات لأكثر من (٢٣٠) ألف جهاز إلكتروني في (١٥٥) دولة حول العالم، حسب وكالة اليوروبول (Europol).

وقد تستخدم الجماعات الإرهابية هذه البرامج في أغراض تخريبية أخرى بخلاف جمع الأموال، وهذا ما ظهر جلياً في هجوم (NotPetya) لعام ٢٠١٧م، حيث تم استهداف الأنظمة الخاصة بأجهزة "الويندوز" في جميع أنحاء العالم، وأغلقت ملفات النظام، إلا أن المهاجمين لم يطالبوا بفدية مالية، وإنما كان الهدف الرئيس هو تدمير البيانات فحسب.

3. هجمات الحرمان من الخدمة (Distributed Denial – Of - Service):

تستخدم التنظيمات الإرهابية هجمات رفض الخدمة (DDoS)، لتشويه مواقع الويب والخوادم، وتعطيل خدمات الشبكة، بهدف استغلال موارد التطبيقات، حيث تتسبب الجهات المنفذة للهجمات في تدفق الموقع بحركات مرور خاطئة، ما يؤدي إلى إضعاف وظائف موقع الويب أو تعطيلها تماماً، وذلك باستخدام الآلاف من الأجهزة، فيما يعرف باسم "شبكات الروبوت" (Joan, 2022).

وفي عام ٢٠١٧م، شن تنظيم داعش أول سلسلة ناجحة من هجمات (DDoS) على شبكة الإنترنت المظلمة، وأطلق عليها "**مدفع الخلافة**"، حيث تم استهداف البنية التحتية العسكرية والاقتصادية والتعليمية، ما أدى إلى تعطل الخدمات عبر الإنترنت.

ويؤكد الباحث على أن التنظيمات الإرهابية لن تتمكن من شن تلك الهجمات وتنفيذ البرامج الضارة، إلا من خلال الانتحال البيومتري للقياسات الحيوية، أي عن طريق محاكاة "الصوت، أو الوجه، أو بصمة اليد، أو العين"، وهو ما يعرف بالتزييف العميق.

المطلب الثاني- التهديدات الاستخباراتية:

تتمكن التنظيمات الإرهابية من استغلال تقنية التزييف العميق في خرق الأنظمة الأكثر أهمية وأماناً في الدولة، والاضطلاع بأعمال التجسس، وذلك بتجاوز فحص الأمان البيومتري ونظم القياسات الحيوية، عن طريق تقديم عينة بيومترية مطابقة تماماً لعينة المستخدم الحقيقي، بغرض الوصول غير المصرح به إلى المرافق الآمنة والمعلومات الحساسة، ما يشكل تهديدات خطيرة بعيدة المدى، لا تعرض خصوصية الأفراد للخطر فحسب، بل تهدد أيضاً المؤسسات والحكومات (جاد الله، ٢٠٢٢).

ومن أبرز البرامج التي تسخرها التنظيمات الإرهابية في أعمال التجسس، هي ما يلي:

1. هجمات الهندسة الاجتماعية:

هي شكل من أشكال التَصَيُّد الاحتيالي[1]، وهي عبارة عن برامج تثبت نفسها عندما يوافق المُستهدَف على شروط وأحكام برنامج يبدو شرعياً، دون قراءة التفاصيل الدقيقة، أو التثبت من صحة مصدره (Rebecca, 2021).

ويتم الاطلاع على كافة البيانات الخاصة بالمُستهدف بمجرد فتحه للمرفقات، أو النقر على الروابط المُضمَّنة (الارتباطات التشعبية) في البريد الإلكتروني، وهناك أنواع من هجمات التَصَيُّد الاحتيالي، ومنها: التَصَيُّد بالرمح، مايكروسوفت ٣٦٥، التصيد عبر وسائل التواصل الاجتماعي، التَصَيُّد الصوتي.

وتعد قرصنة "سولار ويندز" في ٢٠٢٠م، من أبرز الأمثلة التي أثرت في البنية التحتية الحيوية، وهي شكل من أشكال التصيّد الرمحي المُوجَّه، حيث استهدفت العملية اختراق برامج صممتها شركة البرمجيات "سولار ويندز"، ما سمح للمتسللين بالدخول إلى آلاف الشركات والإدارات الحكومية التي تستخدم منتجات الشركة، ومنها: وزارة الدفاع الأمريكية، ووكالة الأمن القومي، ومصلحة الأمن السري (المكلفة بحماية الرئيس الأمريكي)، ووزارة الخزانة الأمريكية، وغير ذلك من الوكالات الحكومية الأميركية، فضلاً عن تضرر بعض المنظمات كـ(الناتو، والبرلمان الأوروبي، ومايكروسوفت، وبعض مراكز الأبحاث)، وغير ذلك.

2. برامج التجسس (Spyware):

هي برامج تعمل بشكل خفي على جهاز الكمبيوتر، فهي تعمل بهدوء في الخلفية، وهذا النوع من البرامج لا يستهدف تعطيل عمليات الجهاز، وإنما يستهدف المعلومات الحساسة، بما في ذلك (المواقع التي يزورها المستخدم، والأشياء التي تقوم بتنزيلها، وأسماء المستخدمين، وكلمات المرور، ومعلومات الدفع، ورسائل البريد الإلكتروني المرسلة والمستقبلة)، ولعل من أخطر أنواع هذه البرامج ما يسمى بـ(keylogger)، وهو برنامج يستطيع تسجيل كل ما يُكَتَّب على لوحة المفاتيح، بمجرد الضغط عليها (Joan, 2022).

3. هجوم الرجل في الوسط (Man-in-the-Middle Attack):

[1] التَصَيُّد الاحتيالي: عبارة عن عمليات احتيالية تعتمد على التفاعل البشري، وتهدف إلى خداع الناس، وتجعلهم يفصحون عن معلومات سرية وشخصية، أو الحصول على أموال.

https://www.jpsa.ac.ae/journal/vol15/iss2/2
DOI: 10.69672/3007-3529.1028

يسهل هذا الهجوم مع استخدام الشخص المُستهدف لشبكات الـ(Wi-Fi) العامة وغير الآمنة، حيث يتم الحصول على جميع المعلومات المتنقلة (الصادرة والواردة) بين جهاز الشخص المُستهدف والشبكة.

ويؤكد الباحث أيضاً على أن التنظيمات الإرهابية لن تتمكن من شن تلك الهجمات الضارة، سواء التصيد الاحتيالي، أو برامج التجسس، أو هجوم الرجل في الوسط، إلا من خلال الانتحال البيومتري للقياسات الحيوية، أي عن طريق محاكاة "الصوت، أو الوجه، أو بصمة اليد، أو العين"، وهو ما يعرف بالتزييف العميق.

المطلب الثالث- التهديدات المجتمعية:

تضطلع التنظيمات الإرهابية بترويج محتويات مزيفة من الوسائط المتعددة، بغرض تمزيق وحدة الصف، وإحداث الفرقة المجتمعية، وتنفيذ الأعمال الإرهابية، ونشر الأفكار المتطرفة، وذلك كما يلي:

أولاً- تمزيق وحدة الصف:

تستطيع التنظيمات الإرهابية تثبيط الهمم والعزائم في صفوف الجيش، وذلك بتداول معلومات مُضلِّلة حول القدرات أو الإمكانيات العسكرية، كعصيان الجنود للأوامر والتعليمات الصادرة لهم من القادة العسكريين، أو هروبهم من ميدان المعركة، وهو ما يسبب الارتباك على المستوى التكتيكي، خاصة أثناء الحروب والصراعات المسلحة، أو إنشاء محتوى يُظهر كبار القادة العسكريين وهم يتلفظون بعبارات عنصرية، أو يعبرون عن ازدرائهم لجنودهم أو رؤسائهم السياسيين، أو يتآمرون ضد الدولة ويخططون لإسقاطها، أو بث مقاطع فيديو لجنود العدو وهم يقتحمون مدينة ما (متنازع بشأنها)، ويسيطرون عليها، ويرفعون شعاراتهم وأعلامهم، وغير ذلك من الأمور التي تتسبب في إحباط المعنويات، وتفكيك الجبهة الداخلية (Lu Jin, 2023).

وقد شاهدنا مؤخراً - في مارس ٢٠٢٢م – فيديو مفبرك للرئيس الأوكراني، وهو يناشد جنوده بعدم مواصلة المعركة ضد روسيا، وضرورة إلقاء أسلحتهم والاستسلام.

ثانياً- الفرقة المجتمعية:

تؤدي الفيديوهات المُختَلَقة وغيرها من محتويات الوسائط المُزيفة إلى تفاقم الانقسامات الاجتماعية، وإحداث التَناحُر بين الأقليات المُضطهَدة، فضلاً عن إثارة الفتن والنعرات القبلية والعرقية، والنزاعات الدينية والطائفية والمذهبية، وهو ما يُفْضِي إلى تمزيق الوحدة الوطنية (Alisdair, 2022).

ومثال ذلك، اضطلاع شاب مسلم بنشر صورٍ أو فيديوهات إباحية لفتاة مسيحية، بغرض الانتقام منها، والتشهير بسمعتها؛ وتزداد الأمور تعقيداً في ظل هيمنة وسائل التواصل الاجتماعي، وسرعة تداولها لمختلف الوسائط المتعددة، وهو ما قد يتسبب في تصادمات طائفية، تؤدي إلى بث الفرقة، وروح الكراهية بين صفوف المجتمع.

ثالثاً- تنفيذ الأعمال الإرهابية:

أدركت التنظيمات الإرهابية أهمية استخدام تقنيات التزييف العميق كأسلحة سرية، حيث تسمح بتنفيذ عمليات إرهابية دون أن يتم اكتشافها، فقد بات من السهل تلفيق الأدلة الكاذبة، أو التلاعب بلقطات المراقبة، أو تزوير صوت ضحاياهم، مما يقوض الثقة في السجلات المرئية، ويعيق العمليات الاستخباراتية.

وتزيد تقنية التزييف العميق من قوة وفعالية الهجمات السيبرانية الإرهابية، حيث تُمكن أعضاءها من تحقيق **الاتصالات السرية**، فالتزييف العميق لا يغير المحتوى اللفظي فحسب، بل يغير أيضاً الخصائص المرئية لكيفية نقل

الرسالة، سواء كان ذلك يشمل حركة فم الشخص وهو يقول شيئاً لم يقله بالفعل، أو سلوك يؤديه ولم يفعله (فارس، ٢٠٢٢).

ويمكن استخدام التزييف العميق لإنشاء هويات أو وثائق سفر جديدة أو تغييرها، بغرض التخفي، وتنفيذ المخططات الإجرامية في سرية تامة، وعلى سبيل المثال، فقد استخدم الأفراد المتهمون في هجمات باريس عام ٢٠١٥م، جوازات سفر مزورة للحصول على قرض مالي قبل تنفيذ هجماتهم (Adriana, 2022).

وفي إطار دعم الجماعات الإرهابية لسرية الاتصالات والمعاملات، فقد ألزمت هيئة تحرير الشام أعضاءها – في مايو ٢٠٢٠م - بالتوقف عن استخدام منصات التواصل الاجتماعي الشهيرة مثل: الماسينجر (Messenger) - والفايبر (Viber)، واستبدلتها بتطبيقات أخرى بديلة، تعتمد على التشفير بين الأطراف، ومنها: وایر (Wire)، وثريما (Threema)، وآي مسج (IMessage)، وداست (Dust).

ولجأت أيضاً إلى استخدام "**العملات الرقمية المشفرة**" مثل "بيتكوين" (Bitcoin)، حيث توفر طبيعتها المشفرة إخفاء الهوية، مما يجعلها وسيلة جذابة لتلك الجماعات، حيث يتم سرقتها من "المحافظ الساخنة"، أو تداولها عبر تقنية الـ(Blockchain)، بغرض شراء المستلزمات اللازمة لتنفيذ الأعمال الإرهابية من خلال شبكة الإنترنت المظلمة، والتي تعمل كسوق للمواد والأسلحة والوثائق المزيفة (David, 2023).

ولاحظ المحققون في تفجيرات سريلانكا عام ٢٠١٩م، أن عدد المعاملات في محافظ البيتكوين التي يستخدمها تنظيم داعش لجمع الأموال زاد بشكل ملحوظ قبل التفجيرات، وكذا الحال بالنسبة لهجمات باريس في عام ٢٠١٥م.

رابعاً- نشر الأفكار المتطرفة:

استخدمت المنظمات الإرهابية في الماضي وسائل التواصل الاجتماعي والويب المظلم؛ لتنسيق الهجمات، ونشر الدعاية، وتجنيد أتباع جدد، وأصبح الإرهاب الإلكتروني يشكل خطورة بالغة، حيث حقق تنظيم داعش – على سبيل المثال - نمواً غير مسبوقٍ في زيادة أعداد المقاتلين الإرهابيين الأجانب الذين سافروا إلى مناطق النزاع، نظراً لشعبية هذه المنصات الاجتماعية وانتشارها (Bryan, 2023).

وأصبح اليوم - في ظل التقنية المتطورة للتزييف العميق - من السهل إنشاء حسابات مزيفة، وانتحال شخصية المستخدمين العاديين على منصات الشبكات الاجتماعية، وهو ما يمكن الجماعات الإرهابية من نشر أيدولوجيتها الغلواء بسهولة أكبر وبخطر أقل، فضلاً عن تسهيل جهود الهندسة الاجتماعية في محاولة للحصول على المعلومات المطلوبة، أو لدعم جهود التطرف.

ويعد روبوت الدردشة (ChatGPT) وغيره من أنظمة الذكاء الاصطناعي المولدة للغات، أدوات هامة تستخدمها الجماعات الإرهابية لخدمة أغراضها المتطرفة، حيث يتم التحايل - باستخدام التزييف العميق - على أنظمة الأمان الخاصة بـ(ChatGPT)، وتجاوز قيود الإشراف على المحتوى؛ لإنتاج محتوى ضار أو التلاعب بالنموذج اللغوي، بغرض إنشاء مقاطع فيديو مزيفة، والترويج للمعلومات المضللة والرسائل التحريضية، والدعاية المفبركة، وحشد الدعم، وتشويه سمعة ممثلي الدولة ومؤسساتها الشرعية، فضلاً عن إمكانية انتحال شخصية أحد القادة السياسيين، أو كبار القادة العسكريين والأمنيين، بغرض جمع المعلومات الهامة والحساسة (Alisdair, 2022).

https://www.jpsa.ac.ae/journal/vol15/iss2/2
DOI: 10.69672/3007-3529.1028

وتستغل الجماعات اليمينية المتطرفة والنازية في الغرب، منصة الـ(ChatGPT)؛ لإنشاء ألعاب فيديو جذابة تتضمن الصور العنيفة، والمؤثرات السمعية والبصرية، والرسوم المتحركة، فضلاً عن أسلحة D3" (تتفاقم المشكلة بسبب سهولة الوصول إلى طابعات D3)، ضمن استراتيجية التجنيد وحشد الدعم، وتستخدم هذه المنصات بأقل تكلفة، حيث تحتاج فقط لاتصال بالإنترنت، وجهاز يمكن من خلاله إصدار تعليمات مصممة بعناية إلى منصة الذكاء الاصطناعي (Adriana, 2022).

وهذا ما أثار مخاوف كل من المدير العام لجهاز MI5 (هيئة المخابرات العسكرية بالمملكة المتحدة)، ومدير مكتب التحقيقات الفيدرالي الأمريكي في قمة استخبارات العيون الخمس[1] في كاليفورنيا، في أكتوبر ٢٠٢٣م، حيث استطاعت التنظيمات المتطرفة سواء كانوا ذئاباً منفردة أو أعضاء في منظمات إرهابية أخرى عنيفة، الاستفادة من البرامج التي تدعم الذكاء الاصطناعي مثل (ChatGPT)؛ لإعداد وتنفيذ أعمال تخريبية، واستشهدا في التدليل على ذلك بتصريحات تنظيم داعش، والتي أعلنت في ديسمبر ٢٠٢٢م على موقع (Rocket.Chat)، وهي منصة اتصالات مفتوحة المصدر، أن التنظيم بدأ في الاعتماد على تقنية (ChatGPT) لتعزيز وحماية الخلافة المتجددة.

المطلب الرابع- تهديدات الأمن القومي:

استغلت التنظيمات الإرهابية تقنيات التزييف العميق لخدمة أغراضها المتطرفة، وباتت تهدد الأمن القومي، وذلك من خلال اضطلاعها بتقليل الثقة في مؤسسات الدولة ورموزها، وتشويه الديمُقراطيّة، وتقويض الصحافة، والإضرار بالمصالح الاقتصادية، وذلك كما يلي:

أولاً- تقليل الثقة في مؤسسات الدولة ورموزها:

بات بإمكان التنظيمات الإرهابية أن تنال من هيبة الدولة ورموزها، وتستهدف مسؤولين أو إدارات حكومية محددة، وذلك بإشاعة مقاطع فيديو دعائية لضابط شرطة يتصرف بعنف، أو قاض يناقش سراً طرق التحايل على النظام القضائي، أو إضرام المواطنين للنيران في الأبنية الحكومية، وتفجير مولدات الكهرباء، وقطع الطرق والسكك الحديدية، فضلاً عن أعمال الشغب، والاعتصامات، والمظاهرات، والإضرابات، والاحتجاجات الفئوية والعمالية، وكل هذه الأمور وغيرها بغرض تَأْجِيج مشاعر الجمهور، وإشاعة الفوضى، وخلخلة الاستقرار الداخلي، وزعزعة الأوضاع السياسية (Sarah, 2022).

وفي هذا الإطار أطلقت مؤخراً جماعة الإخوان المسلمين دعوات لإحداث فوضى في مصر تحت شعار "حراك ١١ نوفمبر"، ودعمت ذلك ببث فديوهات قديمة مفبركة تضمنت الاحتشاد في الميادين، وإحداث اضطرابات عارمة كإحراق المواطنين لبعض المنشآت والمواصلات العامة؛ لحث الجماهير على النزول إلى الميادين والاعتصام لمواجهة الغلاء، والمطالبة بإسقاط الدولة المصرية وإعادة حكمها.

ثانياً- تشويه الديمُقراطيّة:

يتجلى التأثير السلبي لتقدم تكنولوجيا المعلومات والاتصالات وانتشار التزييف العميق، على العملية الديمُقراطيّة ونتائج الانتخابات، حيث بات يسهل تَقْبِيح صورة وسمعة المرشح السياسي، عن طريق بث المعلومات المضللة، وإنتاج

[1] العيون الخمس (بالإنجليزية: Five Eyes)، اختصارا "FVEY"، هو مصطلح يُشير إلى "تحالف استخباراتي"، يشمل كلّ من الولايات المتحدة، والمملكة المتحدة، وكندا، وأستراليا، ونيوزيلندا، بموجب المعاهدة البريطانية الأميركية متعددة الأطراف.

مقاطع فيديو تظهر مرشحاً ينخرط في فعل شائن، أو جنسي، أو يدلي ببيان مثير للجدل بشكل خاص، وهو ما قد يؤثر على الرأي العام بالسلب، ويغير مسار الانتخابات.

وهذا ما فعلته الجهات الضارة - إبان الانتخابات الرئاسية الأمريكية لعام ٢٠١٦م - حيث دفعت بمجموعة من المحتويات الزائفة لتقويض حملة هيلاري كلينتون الرئاسية، واستقطاب المجتمع الأمريكي بشكل عام، وانتشرت تلك المعلومات المزيفة بفضل سرعة وسائل التواصل الاجتماعي.

ثالثاً- الإضرار بالمصالح الاقتصادية:

لم تعد الجهات الضارة في حاجة لإرسال بريد إلكتروني مُزيَّف لإقناع موظف المؤسسة بتحويل الأموال، فقد أصبح من السهل تحقيق ذلك بمكالمة هاتفية تجعل صوت المتصل يحاكي صوت المدير المالي تماماً، وذلك باستخدام تقنيات التزييف العميق، كما يمكن التحايل على آليات المصادقة القائمة على الكاميرا، وعمليات التحقق من الهوية؛ للسماح بولوج المستخدمين غير الشرعيين لموارد وخدمات المنظمة، ومن ثم تسريب ما تحويه من معلومات هامة قد تضرر بالأوضاع الاقتصادية، أو ابتزاز الشركات للحصول على مبالغ مالية مقابل عدم تسريبها.

وعلى سبيل المثال، انتحل شخص في أواخر عام ٢٠٢٣م – ثبت انتماؤه فيما بعد لإحدى التنظيمات الإرهابية – صوت الرئيس التنفيذي لإحدى شركات الطاقة الأمريكية الكبرى، وأجرى اتصالاً بإحدى مسؤولي القطاع المالي بالشركة، وأصدر له تعليمات بتحويل مبلغ (٢٤٣ ألف دولار)، وبالفعل انصاع الموظف للتعليمات، بعدما تأكد من تطابق اللهجة والإيقاع الصوتي للمدير التنفيذي.

وقد يتم الإضرار بجهود التسويق الرقمي، وذلك ببث مقاطع فيديو مُضلِّلة تفيد بأن منتجات شركة ما غير أصلية أو غير مطابقة للمواصفات القياسية، وكذا إعلان مدير شركة ما بأن شركته فقدت جميع المعلومات الخاصة بعملائها، أو أن الشركة على وشك الإفلاس، ما قد يؤدي إلى الإضرار بسمعة العلامة التجارية للشركة، ومن ثم الإضرار بالمصالح الاقتصادية القومية (Fiedelholtz, 2021).

رابعاً- تقويض الصحافة:

تسعى التنظيمات الإرهابية إلى تقويض الصحافة ومصادر المعلومات الجديرة بالثقة، وذلك بترويجها لمحتويات الوسائط المتعددة من صور، وفيديوهات، ورسائل نصية، وصوتية مفبركة، وهو ما يؤدي إلى التشكيك في مصداقية الأخبار والمعلومات التي تقوم ببثها، وبالتالي تنعدم الثقة في المؤسسات الإخبارية البارزة (Yuri, 2019).

ويرى الباحث أن التزييف العميق لا يختلف عن أي تقنية مبتكرة، حيث يمكن استخدامه كأداة لتحسين حياة الناس، وتحقيق النجاحات على المستوى الاستثماري، والاقتصادي، والأمني، والفني، وغير ذلك، بيد أنه يمكن استغلاله من قبل التنظيمات الإرهابية كسلاح فَعّال من أسلحة الحرب النفسية، حيث يسهل اصطناع محتويات رقمية مزيفة، وبثها بسرعة هائلة عبر وسائل التواصل الاجتماعي؛ لإلحاق الضرر بالسمعة، وتلفيق الأدلة، والاحتيال على الجمهور وتخويفهم، وتقويض الثقة في المؤسسات الديمقراطية، فضلاً عن إثارة المشاكل الدينية والسياسية بين مواطني الدولة الواحدة، وخلق شرخ في البناء القومي للمجتمع، وبالأخص في الدول التي تتعدد فيها التركيبات السكانية، وتتكون من أكثر من عنصر وطني، وكل هذا يمكن تحقيقه بأقل الموارد والتكاليف، وعلى نطاق أعم وأشمل.

https://www.jpsa.ac.ae/journal/vol15/iss2/2
DOI: 10.69672/3007-3529.1028

المبحث الثالث

آليات مواجهة هجمات التزييف العميق

تمهيد وتقسيم:

نستعرض في هذا المبحث آليات مواجهة هجمات التزييف العميق، ولعل من أبرزها: التدابير الاحْتِرَازِيَة ضد هجمات التزييف العميق، ثم نبين أهم سبل الحماية الأمنية لمنع التزييف العميق، ثم نوضح المواجهة التشريعية لجرائم التزييف العميق، وذلك كما يلي:

المطلب الأول- التدابير الاحْتِرَازِيَة ضد هجمات التزييف العميق:

نتناول فيما يلي بعض التدابير الاحترازية اللازمة للرصد والاستجابة ضد هجمات التزييف العميق، ومنها: تحصين الدفاعات السلوكية، التحديث المنتظم، القرصنة الأخلاقية، حوكمة الوصول، تدابير الكشف والمنع، وذلك على النحو التالي:

أولاً- تحصين الدفاعات السلوكية:

ينبغي تصميم برامج توعوية تهدف إلى تثقيف الجمهور حول مخاطر التزييف العميق، وتعزيز التفكير النقدي في المواد المتداولة، وتقييم مصداقيتها، والبحث عن مصادرها، فضلاً عن أهمية إِحَاطَتِهم ببعض الأمور الوقائية، والتي تندرج ضمن "**تعليمات ونصائح الأمن السيبراني القياسية**"، ولعل من أبرزها ما يلي:

1. تجنب منح الأُذون للتطبيقات التي تتطلب السماح بالوصول إلى الكاميرا، أو الصور، أو قائمة الأسماء، أو غير ذلك كشرط لكي تعمل؛ لما يشكله من خطورة بالغة على إمكانية انتهاك الخصوصية وخرق البيانات الحساسة (Bharatendra, 2022).
2. تجنب اتصال الأجهزة الذكية بشبكات الواي فاي (Wi-Fi) المجانية والمتاحة بالأماكن العامة، فضلاً عن أهمية استخدم شبكة افتراضية خاصة (VPN)؛ لتمرير البيانات الحساسة والدقيقة من خلالها.
3. تجنب العفوية، وعدم النقر على أي رسائل بريد إلكتروني من جهات غير موثوقة.
4. تجنب استخدم كلمات المرور الشائعة، والعمل على تعقيدها وتَفَرُّدها لمقاومة أنواع الهجوم المتعددة، فضلاً عن أهمية تدوير كلمات المرور المميزة بشكل روتيني.
5. تجنب مشاركة كلمة المرور، وذلك بأن يكون لكل حساب تسجيل دخول وحيد؛ لضمان إشراف واضح، ومسار تدقيق نظيف.

ثانياً- التحديث المنتظم (Regular Updating):

يجب عمل تحديثات لكافة البرامج وأنظمة التشغيل الخاصة بالمنظمة الأمنية بشكل دوري ومنتظم، ويشترط أن يتم التحديث من خلال قنوات آمنة ومشفرة، أي يتم التحقق من سلامة مصدرها قبل تحميلها على شبكة الأجهزة الخاصة بالمنظمات الأمنية، فضلاً عن أهمية التأكد من أن جميع البرامج الخاصة بمكافحة الفيروسات مُحدَّثة، وتعمل بكامل طاقتها في كل الأوقات (De Lima, 2020).

ثالثاً- القرصنة الأخلاقية (Ethical Hacking):

ينبغي أن تُخضِع المنظمة الأمنية كافة برامجها وأنظمة التشغيل لاختبارات دورية، وهو ما يسمى بـ (القرصنة الأخلاقية)، بغرض التأكد من قدرة النظام الرقمي على حماية نفسه ضد أي خروقات، فضلاً عن كشف نقاط الضعف داخل البيئة السيبرانية ومعالجتها قبل اكتشافها، واستغلالها من قبل المحتالين.

رابعاً- حوكمة الوصول (Access Governance):

هو نظام يسمح للمنظمة بالتحكم في وصول العاملين لديها إلى الشبكات والبيانات الحساسة والتطبيقات الهامة، بناء على الدور الذي يؤديه كل فرد، أو بحسب مسؤولياته، أو الصلاحيات والامتيازات المتاحة له داخل المؤسسة (Bharatendra, 2022).

وهذه التقنية تعمل في إطار مبدأ "**الثقة الصفرية**" (ZTNA) اختصاراً لـ(Zero Trust Network Access)، والذي يفرض ضوابط وصول صارمة للغاية على جميع العاملين الذين يطلبون الوصول إلى أصول العمل كـ(الحسابات، والتطبيقات، والأنظمة، والأجهزة)، فيتم منح أحدهم حق الوصول والأذونات التي يحتاجها لأداء دوره فقط (أي في حدود الحد الأدنى المُطلَق والضروري لأداء الأنشطة المكلف بها)، وهذا من شأنه التقليل إلى حد كبير من المخاطر السيبرانية للمؤسسة.

ويتضمن هذا الإجراء ضرورة تفعيل المصادقة الثنائية، كمصادقة افتراضية للدخول إلى الشبكة، فضلاً عن أهمية الأخذ بالمصادقة البيومترية[1]، كعامل إضافي لتأمين الأنظمة والمكونات الحرجة (De Lima, 2020).

خامساً- تدابير الكشف والمنع:

ينبغي أن تتخذ المنظمات كافة التدابير اللازمة لكشف ومنع أي هجمات عدائية سواء على الشبكة، أو الأنظمة، أو البرامج، ولعل من أبرزها ما يلي:

1. جدار الحماية (Firewalls):

ينبغي على المنظمة الأمنية استخدام جدران الحماية، والتي تعد بمثابة حاجز بين الشبكات الداخلية الآمنة (الإنترانت) والخاضعة للرقابة، وتلك الخارجية غير الموثوق بها كالإنترنت، وذلك من أجل مراقبة حركة مرور الشبكة الواردة والصادرة، والسماح لبعض الحركات بالمرور، وحظر بعضها، استناداً إلى مجموعة محددة من قواعد الأمان.

2. أنظمة منع التسلل (Intrusion Prevention Systems):

تعمل أجهزة الـ(Secure IPS) كأنظمة إنذار مبكر لأي تهديدات تهاجم الشبكة، حيث تضطلع برصد ومراقبة حركة الشبكة، وتحديد الأنشطة الخبيثة، ووقف زحف هذه الأنشطة الضارة، وعرقلة حركة مرورها (Bharatendra, 2022).

ويهدف الباحث من خلال طرحه لهذه التدابير الاحترازية اتخاذ كافة ما يلزم من إجراءات لمنع هجمات التزييف العميق الموجهة من قبل الجماعات الإرهابية لتنفيذ أغراضها المتطرفة.

المطلب الثاني- سبل الحماية الأمنية لمنع هجمات التزييف العميق:

[1] المصادقة البيومترية: تعتمد على الخصائص البيولوجية الفريدة الموجودة لدى كل إنسان؛ للتحقق من هويته، مثل: بصمة اليد، الوجه، وغير ذلك.

https://www.jpsa.ac.ae/journal/vol15/iss2/2
DOI: 10.69672/3007-3529.1028

أشرنا من قبل إلى أن إنشاء التزييف العميق يتم عن طريق تبديل مقاطع الفيديو أو الصور أو الصوت مع الهدف، وهو ما دفع الخبراء والمتخصصين لمكافحة هجمات تلك التقنية بالعديد من السبل الحماية الأمنية، ومنها: سلاسل الكتل، الطب الشرعي الرقمي، البصمة الرقمية، منصة سكوبينو التحليلية، العلامات المائية الرقمية المشفرة، وذلك كما يلي:

أولاً- سلاسل الكتل (Block Chain):

هي تقنية لا مركزية، يستطيع المستخدمون من خلالها تخزين بياناتهم عبر الإنترنت دون استخدام الخوادم المركزية، أي لا يتمتع أي شخص أو مجموعة معينة بالسيطرة، وهو ما يزيد من عوامل أمانها، ودقة معلوماتها (Christian, 2022).

وبذلك يمكن لأي فرد استخدام هذه التقنية للمصادقة والتحقق من صحة المحتوى الرقمي سواء كان مستند فيديو أو صوتًا شخصيًا، فهي عبارة عن سلسلة تضم قاعدة بيانات رقمية مشتركة، ويتم إدراج المعلومات إلكترونياً بتسلسل زمني في شكل كتل مُتَّسِقَة، بحيث يصعب أو يستحيل تغيير النظام أو اختراقه، فإذا أراد المتسللون إفساد هذا النظام، فسيتعين عليهم تغيير كل كتلة في السلسلة عبر جميع الإصدارات الموزعة من السلسلة، وهذا أمر مستحيل، وكلما زاد عدد الأشخاص الذين يوقعون الكتل بتوقيعاتهم الإلكترونية، زاد احتمال اعتباره سجلاً أصلياً، وهو ما يضمن أمن سجل البيانات، ويولد الثقة لدى جميع المشاركين (Peter, 2021).

إلا أن هذا الخيار ليس هو الحل الأفضل لمنع هجمات التزييف العميق، نظراً لأنه غير قادر على تخزين كميات كبيرة من البيانات، فهو مناسب بشكل أكبر لتخزين البيانات المُقسَّمة أو المُصنَّفة، والتوقيعات الإلكترونية؛ لذا من الضروري اتخاذ تدابير إضافية لمواجهة ومنع التزييف العميق.

ثانياً- الطب الشرعي الرقمي (Forensic Analysis):

يُعرف الطب الشرعي الرقمي أو ما يسمى بـ"تحليل الطب الشرعي" بأنه: عملية جمع واستخراج الأدلة والبيانات والمعلومات الرقمية في شكلها الأصلي من الأنظمة الرقمية، ومنها: أجهزة الكمبيوتر، الهواتف الذكية، الشبكات، البرامج، قواعد البيانات، وسائط الاتصال، والأنظمة المدمجة كإنترنت الأشياء، المعلومات التي تنتقل عبر الكابلات، ثم إخضاعها لعملية الفحص والتحليل، ويشار للبيانات الرقمية على أنها مزيج من البصمات الرقمية مثل: (الصوت، الرسائل، آثار الشبكة) (Guarnera, 2022).

ويعد الطب الشرعي الرقمي أمراً بالغ الأهمية لتَعَقُّب مسار الأنشطة الرقمية عبر الإنترنت، فضلاً عن قدرته على استعادة واسترجاع كافة البيانات المحذوفة، وهو ما يُمكِّن المحققين من ربط المعلومات الرقمية بالأدلة المادية، واكتشاف الهجمات السيبرانية المُرتكبة، ونظراً لأهمية الأدلة الرقمية، يجب إدارتها بشكل صحيح لمنع تغييرها أو تعديلها.

أدوات الطب الشرعي الرقمية مفتوحة المصدر:

1. **تشريح الجثة (Autopsy):**

هي أداة مفتوحة المصدر بإمكانها استعادة جميع الملفات المحذوفة من الهواتف الذكية ومحركات الأقراص الثابتة، فضلاً عن أنها تضطلع بتحليلها.

2. **و ايرشارك (Wireshark):**

هي أداة برمجية، وتعد من أفضل أدوات الطب الشرعي للشبكات مفتوحة المصدر، حيث تسمح باعتراض البيانات وفك تشفيرها في الوقت الفعلي، فضلاً عن أنها تلتقط وتحلل كل ما يحدث في الشبكة (Frank, 2023).

3. جناح الطب الشرعي الأكسجين (Oxygen Forensic Suite):

هي واحدة من أدوات الطب الشرعي المحمولة مفتوحة المصدر تساعد على الوصول إلى البيانات من الهواتف المحمولة، من دون أي عوائق، حيث تتجاوز كلمة المرور، أو المطالبة بإيماءات شاشة القفل، وغير ذلك (Guarnera, 2022).

4. التقلبات (Volatility):

هو إطار عمل للطب الشرعي يسمح باستخراج المعلومات مباشرة من العمليات التي تعمل على الكمبيوتر، مما يجعله أحد أفضل أدوات التصوير الجنائي والطب الشرعي، كما أنه يتيح استخراج البيانات من ملفات تفريغ الأعطال في Windows، وملفات DLL، ومآخذ الشبكة، واتصال الشبكة نفسه.

5. مستخرج السائبة (Bulk Extractor):

هي أداة برمجية لاستخراج المعلومات، حيث تقوم بمسح الملفات أو الدلائل أو صور القرص، واستخراج البيانات دون تحليل أنظمة الملفات أو هياكل نظام الملفات، فضلاً عن قدرتها على التعامل مع أي تنسيق من الوسائط الرقمية، بما في ذلك بطاقات الكاميرا، ومحركات الأقراص (الثابتة، والضوئية، وذات الحالة الصلبة) (Bharatendra, 2022).

وصدر مؤخراً عدة برامج رقمية مخصصة للكشف عن أي تلاعب بمحتويات الوسائط المتعددة، ومن أهمها الآتي:

أ. **الطب الشرعي للوسائط (MediFor):** يقوم هذا البرنامج بتطوير الخوارزميات؛ لتقييم سلامة الصور ومقاطع الفيديو تلقائياً، وتزويد المحللين بمعلومات مفصلة حول كيفية إنشاء المحتوى المزيف، وذلك برصد الخوارزميات للتناقضات السمعية والبصرية الموجودة في التزييف العميق، بما في ذلك التناقضات في وحدات البكسل (السلامة الرقمية)، وعدم الاتساق مع قوانين الفيزياء (السلامة المادية)، والتناقضات مع مصادر المعلومات الأخرى (السلامة الدلالية) (Noah, 2021).

ب. **الطب الشرعي الدلالي (SemaFor):** يسعى هذا البرنامج إلى تطوير تقنيات دلالية مبتكرة لتحليل الوسائط، وتتضمن هذه التقنيات خوارزميات الكشف الدلالي، والتي ستحدد ما إذا كانت أصول الوسائط قد تم التلاعب بها أم أصلية، فضلاً عن أنها ستحدد ما إذا كانت الوسائط صادرة عن مؤسسة أو فرد معين.

ويعاب على أسلوب الطب الشرعي الرقمي أنه إجراء لاحق، فهو نوع من "**المصادقة السلبية**"، حيث يتم إجراؤه للتَثَبُّت والتحقق عما إذا كان قد تم العبث بالمحتوى الرقمي، فهو لم يتضمن إشارة مصادقة مُسبَقة إلى الوسائط الرقمية، تُمكِّنه من منع التلاعب أو التزييف بالوسائط الإلكترونية.

ثالثاً- البصمة الرقمية (Digital Footprint):

البصمة الرقمية أو الظل الإلكتروني هي تَعَقُّب مسار البيانات المُخَلَّفَة لشخص ما على النظام، مثل: النشر، أو المشاركة على مواقع الشبكات الاجتماعية، أو المنتديات عبر الإنترنت، أو إرسال واستقبال رسائل البريد الإلكتروني، أو الاشتراك في رسائل إخبارية، أو الموافقة على قبول ملفات تعريف الارتباط على متصفح الويب، أو التسوق عبر الإنترنت، أو ترك مراجعة عبر الإنترنت، وأيضاً إذا قام المستخدم بتسجيل الدخول إلى موقع ويب من خلال اسم مستخدم أو ملف

https://www.jpsa.ac.ae/journal/vol15/iss2/2
DOI: 10.69672/3007-3529.1028

تعريف مسجل، وهذه جميعها تعد آثاراً أو بصمات رقمية يستطيع محللو المعلومات من خلالها تتبع أنشطة الشخص المُستهدف وأجهزته عبر الإنترنت، والحصول على البيانات المهمة، ويمكن الاستعانة في ذلك بتطبيق جوجل (Google Analytics).

رابعاً- منصة سكوبينو التحليلية (Skopino Analytical Platform):

يعد سكوبينو (Skopenow) محرك بحث تحليلي، يستخدم وسائل التواصل الاجتماعي وبيانات الويب مفتوحة المصدر بشكل أكثر كفاءة وفعالية، حيث يقوم (Skopenow) تلقائياً بجمع البيانات من جميع أنحاء الإنترنت، ثم يقوم بفرز وتحليل سريع للبيانات المهيكلة (هي بيانات منظمة، وتستخدم ملفات Excel، أو قواعد بيانات SQL)، وغير المهيكلة (لا تتبع تنسيقًا محددًا مسبقًا)، وعرضها في لوحة معلومات واحدة، فضلاً عن قيامه بتحليل الروابط، حيث يسهم في مساعدة المؤسسات على تصور وتقييم العلاقات والروابط المعقدة بين الكيانات، فغالباً ما يترك المجرمون وراءهم آثاراً للنشاط الرقمي، عندما يقومون بأنشطة غير مشروعة عبر الإنترنت، وغالبا ما تكون هذه الآثار مجزأة، أي غير متصلة في البداية، وهنا يتجلى دور سكوبينو لربط الجناة بالأنشطة غير المشروعة، وإنشاء تصور للعلاقات بين الأطراف المعنية، مما يساعد في تحديد نقاط الضعف، والتهديدات والمخاطر المحتملة، ومراقبتها والتخفيف من حدتها، مع التركيز القوي على أمن المعلومات (Bharatendra, 2022).

ويتضح مما سبق أن كلٌّ من سلاسل الكتل، والبصمة الرقمية، ومنصة سكوبينو التحليلية، والطب الشرعي الرقمي ليست بالتقنيات الفَعَّالَةٌ في الوقت الحالي لمنع هجمات التزييف العميق، نظراً لكونها إجراءات لاحقة تضطلع بالبحث والتقصي عن الانتهاكات المرتكبة، فضلاً عن التحسينات والتطورات المستمرة للشبكات التوليدية (GANs)، وقدرتها على إنشاء مقاطع فيديو مُزيَّفة بدقة عالية لا يمكن اكتشافها بسهولة، ومن ثم، فإن العلاج الحقيقي هو الوقاية من التزييف العميق، وليس مجرد الكشف.

خامساً- العلامات المائية الرقمية المشفرة:

(Encrypted Digital Watermarks):

تستخدم هذه التقنية لحماية المحتوى الرقمي، ويطلق عليها "**المصادقة النشطة**" تمييزاً لها عن تلك السلبية التي تطلق على "تحليل الطب الشرعي" وغيرها من تقنيات التَتَبُّع الرقمية.

والعلامات المائية الرقمية هي إشارات رقمية يتم دمجها ضمن الوسائط المتعددة (النص، الصوت، الرسومات، الصور المتحركة، الفيديو، التطبيقات التفاعلية)، وتتم هذه العملية في وقت التقاط الفيديو أو قبل توزيعه، وتظهر هذه العلامات المائية في مسارات الأصوات أو إطارات الفيديو الأصلية، وتختفي عند التزييف كتبديل الوجوه Sebastian, (2022).

ويجب إخفاء إشارة المصادقة الرقمية بما تتضمنه من البيانات الوصفية التعريفية (المُضمَّنة بالعلامة المائية) داخل الوسائط المتعددة، أي عدم إدراكها بالعين المجردة، ويكون ذلك بتشفيرها أو ترميزها بشكل مستمر وغير محسوس، ثم تضطلع البرامج أو الأجهزة - التي تدعم وحدة فك ترميز العلامة المائية - باستخراج العلامة المائية من إشارة الوسائط، والوصول إلى البيانات الوصفية (Joan, 2022).

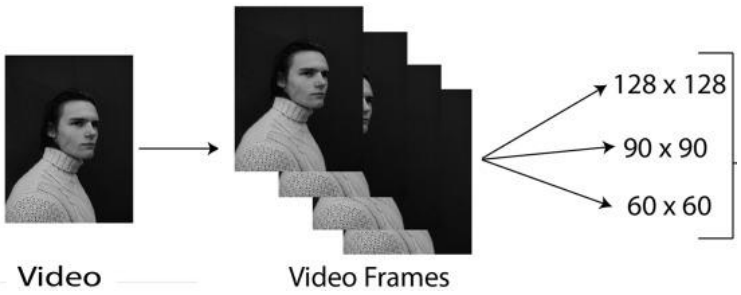
وفي إطار حماية وتأمين الوسائط الرقمية، يجب دمج خوارزميات التشفير مع تقنية سلاسل الكتل (Blockchain)، والاستفادة من طبيعة الأخيرة غير القابلة للتغيير، ويتحقق ذلك بتسجيل البيانات الوصفية المشفرة، وقيمة المجموع الاختباري، على أنها "كتلة" من البيانات في نظام الـ(Blockchain)، وذلك قبل توزيع أي جزء من المحتوى، وبما أن كل كتلة متصلة، واحدة تلو الأخرى، فيصبح من الصعب العبث بأي من البيانات (Joan, 2022).

وتمر عملية تضمين العلامة المائية داخل إطارات الفيديو، بمرحلة تمهيدية، وأربع مراحل رئيسة، وذلك على النحو التالي:

المرحلة التمهيدية- المعالجة المُسبَقة للبيانات (Data Preprocessing):

يتم في هذه المرحلة التمهيدية تحويل مقاطع الفيديو إلى إطارات، ثم يتم تمرير جميع إطارات الفيديو كمصفوفة في الشبكة العصبية التلافيفية ثلاثية الأبعاد[1](3D CNN).

وفي هذه المرحلة أيضاً يتم تغيير حجم الإطارات من الحجم الأصلي إلى أحجام مختلفة، مثل: (١٢٨ × ١٢٨) - (٦٠ × ٦٠) - (٩٠ × ٩٠)، كما هو موضح بالشكل التالي:



وتعد هذه المرحلة التمهيدية بالغة الأهمية، نظراً لأنه كلما زاد حجم الإطار، أصبح التدريب أكثر فعالية، فضلاً عن التحقق من سلامة النهج المُقْتَرَح عن طريق الاختبار على أحجام المدخلات المختلفة (De Lima, 2020).

المرحلة الأولى- توليد نموذج الميزات (Generate Model Features):

هذه المرحلة يتم من خلالها تغذية الشبكة العصبية التلافيفية ثلاثية الأبعاد (3D CNN) بمجموعة بيانات إدخال الفيديو (التي تمت معالجتها مسبقًا)؛ لإنشاء وتوليد **"نموذج الميزات"**، ويعرف الأخير بأنه: عبارة عن شبكة الميزات المُدَرَّبة على اكتشاف المشاهد المختلفة، حيث تضطلع بتحديد (طول، وعرض، وارتفاع، وبُعْدُ البيانات)، وغير ذلك من الميزات (Frank, 2023).

المرحلة الثانية- تشفير العلامة المائية (Watermark Encryption):

هذه المرحلة يتم من خلالها ترميز العلامة المائية، وتضمينها في إطارات الفيديو لمنع إعادة تبديل نفس العلامة المائية، ويتم ذلك بتمرير كل من نموذج الميزات (الذي تم إنشاؤه في المرحلة السابقة)، والعلامة المائية إلى شبكة الخصومة التوليدية (GAN) للتدريب؛ لمطابقة الأول لإطارات الفيديو، وغرس الثانية ومطابقتها أيضاً لإطارات الفيديو، وبالتالي تصبح العلامة المائية غير مرئية (Guarnera, 2022).

[1] **الشبكة العصبية التلافيفية (CNN):** هي فئة من الشبكات العصبية الاصطناعية، تُستخدم لتحليل الصور المرئية، والمهام التي تنطوي على معالجة بيانات البكسل، نظراً لأنها تعتمد على عملية رياضية تسمى الالتفاف بدلاً من ضرب المصفوفة العامة في طبقة واحدة على الأقل من طبقاتها.

https://www.jpsa.ac.ae/journal/vol15/iss2/2
DOI: 10.69672/3007-3529.1028

وتحمي هذه المرحلة المعلومات داخل الفيديو بتشويش محدد، حيث يتم تضمين العلامة المائية المشفرة مرة واحدة في إطارات الفيديو، ولا يمكن الوصول إليها وقراءتها إلا إذا كان "نموذج الميزات" معلوم.

المرحلة الثالثة- فك تشفير العلامة المائية (Decryption of Watermark):

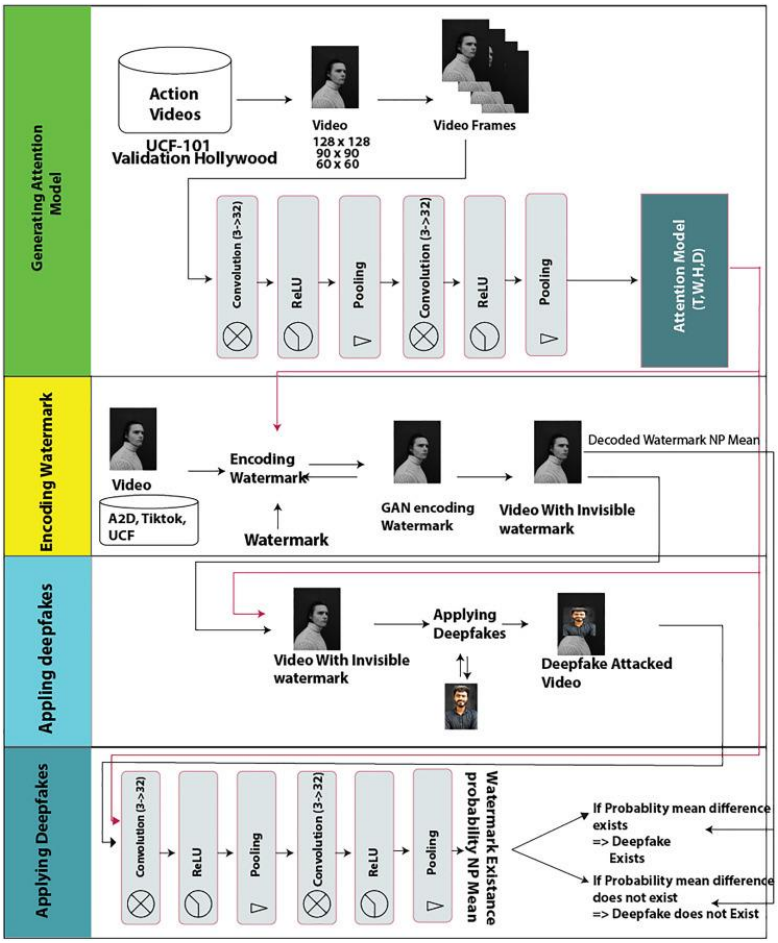
تكون وظيفة هذه المرحلة التحقق من صحة الفيديو باستخدام خوارزمية استخراج العلامة المائية، وتتطلب الأخيرة توفير نفس "نموذج الميزات" لفك تشفير العلامة المائية قبل أن يتمكن المهاجم من تطبيق التزييف العميق على الإطار (De Lima, 2020).

وذلك يعني أن المهاجم لن يتمكن من إنشاء تقنية التزييف العميق ما لم يكن لديه "نموذج الميزات" الناتج عن النهج المُقْتَرَح، فغياب "نموذج الميزات" يجعل هجوم التزييف العميق مستحيلاً.

المرحلة الرابعة- التحقق من التزييف العميق (Deepfake Verification):

يأتي دور هذه المرحلة بعد تَوَفُّر "نموذج الميزات" المطلوب؛ ليتم التحقق من وجود العلامة المائية باستخدام درجة احتمال التواجد، أي أن احتمال وجود العلامة المائية يضمن عدم تزييف الفيديو (Noah, 2021).

ويوضح الشكل التالي شرح المراحل الأربعة سالفة الذكر، حيث يتألف النموذج التدريبي من الشبكة العصبية التلافيفية (CNN)، فضلاً عن شبكتين عصبيتين عميقتين مختلفتين، أحداهما: للتشفير، والأخرى: لفك التشفير، وذلك كما يلي:



مما سبق يتضح أن تقنية العلامات المائية الرقمية المشفرة يمكنها تحقيق الوقاية من هجمات التزييف العميق، نظراً لأنها "**إجراء استباقي**" يتم إعداده وترميزه قبل أي هجمات احتمالية، وهذا ما يدعونا لأن نتناول المراحل التدريبية التي تمر بها العلامة المائية الرقمية.

ويرى الباحث إجمالاً لما سبق، أنه في ظل التقدم المستمر في النماذج التوليدية، لن تتحقق الحماية الكافية ضد هجمات التزييف العميق باكتشافها وتتبعها، بل يجب العمل على المنع والوقاية ضد تداعياتها المتطورة والمتلاحقة، وهذا ما أبرزناه من خلال تقديم نهج منعي يضطلع برصد الهجمات ومنعها، حيث يتم تضمين العلامة المائية المخفية في ميزات إطار الفيديو، حتى يُجبَّر المتسللون على البحث عن نموذج الميزات، من أجل فك تشفير العلامات المائية، وتطبيق التزييف العميق.

المطلب الثالث- المواجهة التشريعية لجرائم التزييف العميق:

نستعرض من خلال هذا المطلب سبل المواجهة التشريعية لجرائم التزييف العميق على الصعيدين الدولي والإقليمي، وذلك بتناول التشريعات الخاصة بحماية بعض دول المنطقة العربية للمحتوى الرقمي من التزييف والتلاعب، وتحقيق قدرة الدول على توفير بيئة سيبرانية آمنة، فضلاً عن الوقوف على التشريع الخاص بالاتحاد الأوروبي، والقانون الأمريكي لحماية حق المؤلف، وذلك على النحو التالي:

أولاً- على الصعيد الإقليمي:

نتناول فيما يلي بعض التشريعات الخاصة بحماية المحتوى الرقمي ضد التزييف والتلاعب، وذلك ببعض دول المنطقة العربية، ومنها: المنظومة التشريعية للإمارات العربية المتحدة، وقانون مكافحة جرائم تقنية المعلومات لسلطنة عُمان، وتشريع المملكة العربية السعودية الخاص بمكافحة الجرائم السيبرانية، وذلك على النحو التالي:

أ. المنظومة التشريعية للإمارات العربية المتحدة بشأن جرائم التزييف العميق:

يستعرض الباحث فيما يلي القوانين المتعلقة بالأنشطة المنفذة على شبكة الإنترنت، والتي تنتهجها دولة الإمارات العربية المتحدة في مواجهة الجرائم الإلكترونية، وتداول البيانات، وآليات استخدام الإنترنت، وسياسة النفاذ إليه.

أرسى **القانون رقم (٣٤) لسنة ٢٠٢١م، بشأن الشائعات والجرائم الإلكترونية**[1]، عقوبات سالبة للحرية تتمثل في الحبس والسجن المؤقت، فضلاً عن فرض الغرامات، لكل من يرتكب انتهاكات إلكترونية، قد تمس الدولة أو الغير، حيث نص على تجريم بعض الأفعال، ولعل من أبرزها: نشر معلومات كاذبة، أو معلومات تضر بمصالح وأمن دولة الإمارات، واختراق نظم المعلومات والبيانات الحكومية أو مهاجمتها أو العبث بها، وإنشاء أو تعديل روبوتات إلكترونية لنقل بيانات زائفة في الدولة، وتزوير المستندات الإلكترونية، والاعتداء على البيانات والمعلومات الشخصية، وغير ذلك.

ويعد هذا القانون بمثابة السياج لحماية الأمن الوطني الإماراتي، حيث غلَّظ العقوبات في حالات الإضرار بأنظمة المعلومات بالجهات المصرفية، أو الإعلامية، أو الصحية والعلمية، وكذا بالنسبة لمؤسسات الدولة والمرافق الحيوية؛ لتصل إلى السجن المؤقت وغرامة لا تقل عن (خمسمائة ألف درهم) ولا تزيد عن (ثلاثة ملايين درهم).

[1] مرسوم بقانون اتحادي رقم (٣٤) لسنة ٢٠٢١م، بشأن الشائعات والجرائم الإلكترونية، متاح على الرابط التالي:
https://laws.uaecabinet.ae/ar/materials/law/1526

https://www.jpsa.ac.ae/journal/vol15/iss2/2
DOI: 10.69672/3007-3529.1028

وحرصت أيضاً دولة الإمارات العربية المتحدة على الحفاظ على خصوصية بيانات المتعاملين معها، فقد وضعت **القانون رقم (٢٦) لسنة ٢٠١٥م، بشأن تنظيم نشر وتبادل البيانات في إمارة دبي**[1]، حيث نصت المادة (٩) على إلزام مزودي البيانات بعدم المساس بحقوق الملكية الفكرية، ونصت أيضاً المادة (١٣) على إلزامهم باتخاذ كافة الإجراءات اللازمة للحفاظ على سرية وخصوصية بيانات المتعاملين.

وتفعيلاً لهذه القوانين، وحرصاً على حماية خصوصية المستخدم، فقد تم حصر فئات المحتوى المحظور في (١٩) فئة، وإلزام مالكي مواقع الويب التي تخدم المنطقة والإمارات العربية المتحدة، بضرورة التوافق والالتزام بالتوجيهات الإرشادية لاستخدام الإنترنت (ومن أبرزها: التأكد من سلامة المحتوى الإلكتروني، وعدم إدراجه ضمن الفئات المحظورة)، ضماناً لاستمرارية تقديمهم للخدمات الإلكترونية، وتجنب حجبها.

وتنفيذاً لذلك، تم تمكّين هيئة تنظيم الاتصالات وحدها دون غيرها (بصفتها صاحبة الاختصاص الأصيل)، من إصدار قرار بحجب النفاذ إلى المحتوى المحظور، أو رفع الحجب عن النفاذ إلى محتوى الإنترنت، وذلك وفقاً لتقديرها المطلق والمرتكز على هذه السياسة، وفي حالة الحجب، يتوجب على المرخص لهم حجب النفاذ إلى المحتوى المحظور، مع مراعاة أن عملية الحجب لا تؤثر سلباً على استقرار شبكة وخدمة الإنترنت في الدولة.

ب. موقف المشرع العُماني من جرائم التزييف العميق:

تناول المشرع العُماني أغلب أنواع الجرائم المعلوماتية بالتجريم في **قانون مكافحة جرائم تقنية المعلومات لسلطنة عُمان لعام ٢٠١١م**[2]، حيث نص الفصل الثاني من هذا القانون على عقوبات بالسجن والغرامة؛ لمن يتعدى على سلامة وسرية وتوافر البيانات والمعلومات الإلكترونية والنظم المعلوماتية، وتضمن الفصل الثالث عقوبات سالبة للحرية وغرامات مالية؛ لمن استخدم وسائل تقنية المعلومات لأغراض ارتكاب جرائم معلوماتية، وعاقب الفصل الرابع على حالات التزوير والاحتيال المعلوماتي، وشدد الفصل الخامس بعقوبات وغرامات مالية لمن استخدم الشبكة المعلوماتية للإخلال بالآداب العامة، أو المساس بالقيم الدينية أو النظام العام، أو التعدي على حق محمي قانونا لمؤلف، وشمل الفصل السادس البطاقات المالية بالحماية، حيث عاقب بالسجن والغرامة المالية لكل من زور بطاقة مالية بأية وسيلة كانت أو اصطنع أو صنع أجهزة أو مواد تساعد على ذلك.

ج. موقف المملكة العربية السعودية من جرائم التزييف العميق:

أصدر مجلس الوزراء السعودي القرار رقم (٧٩) الصادر في ٧ / ٣ / ١٤٢٨هـ؛ لمكافحة الجرائم المعلوماتية[3]، وتضمن عقوبات صارمة بالسجن وغرامات مالية عند المساس بالحياة الخاصة، أو التصنت، أو التشهير، وشمل حماية خاصة للبيانات البنكية أو الائتمانية، وجرم أي فعل يحاول إيقاف الشبكة المعلوماتية عن العمل، أو تعطيلها، فضلاً عن حمايته للقيم الدينية والنظام العام، والآداب العامة، كما شدد على إنشاء منظمات إرهابية أو الانضمام لعصابات منظمة؛ للحصول على بيانات تمس الأمن الداخلي أو الخارجي للدولة، أو اقتصادها الوطني.

---

[1] مرسوم بقانون اتحادي رقم (٣٤) لسنة ٢٠٢١م، بشأن تنظيم نشر وتبادل البيانات في إمارة دبي، متاح على الرابط التالي:
https://laws.uaecabinet.ae/ar/materials/law/1526

[2] قانون مكافحة جرائم تقنية المعلومات لسلطنة عُمان لعام ٢٠١١م، متاح على الرابط التالي:
https://qanoon.om/p/2011/rd2011012/

[3] القرار رقم (٧٩) الصادر في ٧ / ٣ / ١٤٢٨هـ؛ لمكافحة الجرائم المعلوماتية، متاح على الرابط التالي:
https://laws.boe.gov.sa/BoeLaws/Laws/LawDetails/25df73d6-0f49-4dc5-b010-a9a700f2ec1d/1

ثانياً- على الصعيد الدولي:

نتناول فيما يلي بعض التشريعات الخاصة بحماية المحتوى الرقمي ضد التزييف والتلاعب، وذلك بالوقوف على التشريع الخاص بالاتحاد الأوروبي، والقانون الأمريكي لحماية حق المؤلف، وذلك على النحو التالي:

أ. قانون الخدمات الرقمية (The Digital Services Act (DSA:

صدر قانون الخدمات الرقمية (DSA) عن الاتحاد الأوروبي في نوفمبر ٢٠٢٢م، وهو مجموعة شاملة من القواعد الجديدة التي تنظم مسؤوليات الخدمات الرقمية، وتشير الأخيرة إلى الخدمات الوسيطة مثل: مقدمي الخدمات المضيفة، والأسواق عبر الإنترنت، وشبكات التواصل الاجتماعي، ومنصات مشاركة المحتوى، ومتاجر التطبيقات، ومنصات السفر والإقامة عبر الإنترنت.

ويهدف هذا القانون إلى خلق بيئة أكثر أماناً وعدلاً على الإنترنت لجميع المستخدمين في الاتحاد الأوروبي على قدم المساواة، وذلك بحظر الأنشطة غير القانونية والضارة عبر الإنترنت، وانتشار المعلومات المضللة، ومكافحة خطاب الكراهية والتشجيع على الإرهاب، فضلاً عن حماية الحقوق الأساسية في الفضاء الرقمي، من خلال تيسير آليات مجانية للشكاوى والإبلاغ عن أي محتوى أو سلع أو خدمات غير قانونية.

وبموجب هذا النظام، تكون المنصات الرقمية مسؤولة عن الخوارزميات الخاصة بها، ومسؤولة أيضاً عن تقييم المخاطر، واتخاذ كافة التدابير الفعالة للحد من هذه المخاطر، وحماية المستخدمين.

ب. قانون الألفية الرقمية لحق المؤلف Digital Millennium Copyright Law:

أقر الكونغرس الأمريكي قانون الألفية الرقمية لحق المؤلف (DMCA) في عام ١٩٩٨م، وحرص على توفير الحماية الكافية لحقوق الملكية الفكرية، حيث أدانت المادة (١٢٠١) من ذات القانون نوعين من الأنشطة، وهما كالآتي:

1. التحايل على تدابير الحماية التكنولوجية التي يستخدمها مالكو حقوق الطبع والنشر؛ للتحكم في الوصول إلى أعمالهم ومصنفاتهم عبر خدمات البث، وعلى سبيل المثال، من يحاول الوصول بطرائق غير مصرح بها، إلى محتوى العمل أو المصنف المحمي، كاختراق كلمات المرور، أو التحايل على التشفير.
2. تصنيع أو الاتجار في التكنولوجيا المصممة؛ للتحايل على التدابير التي تتحكم في الوصول إلى هذه المصنفات، أو تحمي حقوق مالكي حقوق الطبع والنشر (Ignacio,2020).

وفي عام ٢٠٢٠م، أدخل الكونغرس تعديلاً للقانون الذي نحن بصدده، يتلخص في إنشائه لـ"مجلس مطالبات حقوق الطبع والنشر" (CCB)، وهو يعد بمثابة محكمة اختيارية بديلة للمحكمة الفيدرالية، تختص بالفصل في نزاعات حقوق الطبع والنشر التي تنطوي على مطالبات تسعى للحصول على تعويضات تصل إلى ٣٠ ألف دولار، وهي مصممة لتكون أقل تكلفة، وأسرع من رفع القضايا بالمحكمة الفيدرالية.

وتعد هذه المحكمة البديلة (CCB) نظامًا إلكترونيًا لإدارة القضايا، فهي متاحة لأي شخص، مع أو بدون محام، حيث يقدم المدعي (الطرف الذي لديه مطالبات بحقوق الطبع والنشر) وثائقه ومعلوماته، التي تؤكد حجته في المطالبة بحقوقه عبر الإنترنت، ويتم الفصل فيها عن طريق "جلسات استماع عن بعد" من خلال مؤتمرات الفيديو (Ignacio,2020).

https://www.jpsa.ac.ae/journal/vol15/iss2/2
DOI: 10.69672/3007-3529.1028

الخاتمة

في ختام عرضنا السابق لموضوع "**استخدام التنظيمات الإرهابية للتزييف العميق وسبل المواجهة**"، يمكننا أن ننتهي إلى أن التلاعب بالصور أو مقاطع الفيديو ليست بالتقنية المُستحدَثة، إلا أنه مع الاعتماد على الذكاء الاصطناعي في شتى نواحي الحياة، وشُيوع وسائل الإعلام الرقمي الجديد، فقد ازدادت هجمات التزييف العميق بشكل واضح، وأصبح بثها وترويجها أكثر سهولة ويسر، وباتت شبكة الإنترنت ساحة خصبة لبث تلك الهجمات بما تحمله من مَضارّ على الفرد والمجتمع، وعلى أمن واستقرار الدولة.

وتطرق الباحث في بحثه إلى المفهوم العلمي للتزييف العميق، ثم تناول أهم تصنيفاته المختلفة، وهي: النص، الصورة، الصوت، مقاطع الفيديو، ثم بَيَّنَ أبرز الأساليب المستخدمة لتوليد التزييف العميق، وكيفية إنشاء فيديوهات مبادلة الوجوه والتلاعب بتعابير الوجه البشري، وكان ذلك بالاعتماد على التشفير التلقائي (Autoencoder)، وشبكات الخصومة التوليدية (GAN).

وسلط الباحث الضوء على مخاطر استخدام التنظيمات الإرهابية للتزييف العميق، والتي تمتد تداعياتها وآثارها الضارة إلى ما هو أبعد من الإضرار بالسمعة، حيث تتجلى في استهداف الأمن القومي، والتهديدات الاستخباراتية، والسيبرانية، والمجتمعية، وغير ذلك.

وأوضح الباحث التدابير الاحْتِرَازِيَة اللازم اتخاذها لمواجهة هجمات التزييف العميق، والتي استعرض من خلالها أهم بروتوكولات الأمن السيبراني للرصد والاستجابة، وأفضل الأساليب التقنية لمنع التزييف العميق، بالإضافة إلى الطرائق العلمية لكيفية تضمين العلامة المائية الرقمية.

أولاً- نتائج البحث:

1. وضَّحَ الباحث أن انتشار تطبيقات الذكاء الاصطناعي والتعلم الآلي، ومجانية الحصول على أغلبها، وسرعة بثها للوسائط الإلكترونية من صور ومقاطع فيديو، أَتَاحَ للتنظيمات الإرهابية تطويع تلك التقنيات في الأعمال المتطرفة وغير المشروعة.
2. بَيَّنَ الباحث أن هجمات التزييف العميق لن تقتصر آثارها التخريبية على مستوى الأفراد والجماعات، وإنما تمتد لتزعزع استقرار الدول، وتشيع الفوضى، وتثير البلبلة بين مختلف الشعوب والحكومات، فالعالم اليوم أصبح قرية كونية صغيرة بفضل سيطرة التكنولوجيا الرقمية.
3. دَلَّلَ الباحث على أن ظاهرة التلاعب بالمعلومات لأغراض خادعة، أو ما تسمى بالتزييف العميق، تعد من أحدث الأسلحة المستخدمة في حرب المعلومات العصرية، والتي توصف أيضاً بأنها حرب نفسية.
4. أظهر الباحث أن التنظيمات الإرهابية بإمكانها استغلال تقنية التزييف العميق في خرق الأنظمة الأكثر أهمية وأماناً في الدولة، والاضطلاع بأعمال التجسس، نظراً لقدرتها على تجاوز فحص الأمان البيومتري.
5. أوضح الباحث أنه يسهل للتنظيمات الإرهابية أن تُحدث انقسامات في صفوف الجيش، وتُدمِّر المعنويات، وتفكك الجبهة الداخلية، فضلاً عن بث الفرقة، وروح الكراهية بين صفوف المجتمع، وذلك باستغلالها السيئ لتقنيات التزييف العميق.

6. أَكَّدَ الباحث أن بث الجهات الضارة لمقاطع الفيديو والصور المُضلِّلة، من شأنه الإضرار بسمعة العلامات التجارية للشركات، وهو ما ينعكس سلباً على المصالح الاقتصادية للدول.
7. أَثْبَتَ الباحث أن شبكات الخصومة التوليدية (GANs) تعد من أفضل الأدوات التوليدية للتزييف العميق، وذلك بفضل قدرتها على تحقيق نتائج واقعية عالية الجودة في تراكيب الصور، وتوليف الفيديو، وإنتاج الموسيقى، والمهام الأخرى.
8. برهنَ الباحث أن تضمين العلامات المائية الرقمية المشفرة يعد من أفضل التقنيات لمنع التلاعب بالوسائط الإلكترونية؛ لكونه إجراءً استباقيًا يتم إعداده وترميزه قبل أي هجمات احتمالية.

ثانياً- توصيات البحث:

يطيب لي في ختام تناولي لهذا البحث، أن أشير – على سبيل التوصية - إلى بعض النقاط التي أرى أنها فاعلة ومؤثرة في مجابهة ظاهرة التزييف العميق، والحد من انتشارها، وأوجز تلك التوصيات فيما يلي:

1. تعزيز التعاون الرقمي "أمنياً وقضائياً ومعرفياً" بين أجهزة إنفاذ القانون، والحكومات، والمنظمات الدولية، والخبراء من شركات الأمن السيبراني؛ وإنشاء قنوات اتصال واضحة للاستجابة الفورية للتهديدات والحوادث المتعلقة بالتزييف العميق، والحد من انتشار الجرائم السيبرانية.
2. ضرورة توفير الحماية القانونية الكافية للأفراد والمؤسسات المتضررة جراء هجمات التزييف العميق، فضلاً عن أهمية تعويضهم مادياً وأدبياً بما يتناسب مع الجُرم المُرتكَب حيالهم.
3. حظر مواقع الويب والتطبيقات الرقمية، حال تعمد مالكيها نشر أي محتويات تدعم الفكر والنشاط الإرهابي المتطرف على منصاتها، أو بث أي وسائط مزيفة، فضلاً عن توقيع عقوبات جنائية رادعة.
4. تطبيق نظام "الشرطة المجتمعية على الإنترنت"، وذلك بمشاركة الشرطة مع المجتمع المدني عبر وسائل التواصل الاجتماعي، وهو نهج استباقي لمنع الجريمة، حيث يقيم علاقات مع ضحايا الجرائم الإلكترونية المحتملين، ويعزز المعلومات المشاركة، والتغلب على المخاوف المرتبطة عادة بالإبلاغ عن الجرائم الإلكترونية.
5. النهوض بالوعي والثقافة المجتمعية حول مخاطر وتهديدات التزييف العميق على أمن وسلامة المجتمع ومقدراته المادية والمعنوية، والإِحَاطة ببعض الأمور الوقائية، والتي تندرج ضمن "تعليمات ونصائح الأمن السيبراني القياسية"، والتي من أبرزها ما يلي:
   أ. القرصنة الأخلاقية.
   ب. حوكمة الوصول.
   ج. مبدأ "الثقة الصفرية".
   د. المصادقة البيومترية.

https://www.jpsa.ac.ae/journal/vol15/iss2/2
DOI: 10.69672/3007-3529.1028

6. الحفاظ على النظافة الرقمية، وذلك بالتحديثات المنتظمة لأنظمة التشغيل والبرامج، والمسح الشامل للبرامج التي تخص الطرف الثالث، أو الأجهزة التي هي جزء من سلسلة التوريد، فضلاً عن نشر أنظمة كشف ومنع التسلل (IDPS)؛ لرصد نشاط البرامج الضارة، ووقف زحفها.
7. اِعْتَمَادَ المصادقة النشطة، وذلك بدمج العلامة المائية المشفرة ضمن الوسائط المتعددة، ويتم تشفيرها بشكل مستمر وغير محسوس؛ لتظهر هذه العلامات المائية في مسارات الأصوات أو إطارات الفيديو الأصلية، وتختفي عند التزييف كتبديل الوجوه.

المصادروالمراجع العلمية

Scientific sources and references

First- Arabic References:

Jadallah, Abdulaziz Lutfi, (2022). Cybercrime and Information Security Protection, Egypt: Knowledge Foundation for Publishing.

Fares Al-Amarat (2022). Cybersecurity (concept and challenges of the times), Jordan: Gulf Publishing House.

Second- Websites:

https://laws.uaecabinet.ae/ar/materials/law/1526

https://qanoon.om/p/2011/rd2011012/

https://laws.boe.gov.sa/BoeLaws/Laws/LawDetails/25df73d6-0f49-4dc5-b010-a9a700f2ec1d/1

Thirdly- Foreign References:

Adriana Burlea-Schiopoiu ،Ahmed Jabbar Obaid ،Ghassan H. Abdul-Majeed ، Parul Aggarwal (2022). Handbook of Research on Advanced Practical Approaches to Deepfake Detection and Applications, IGI Global, USA.

Alisdair A. Gillespie (2022). Cybercrime: Key Issues and Debates, Routledge, New York.

Bharatendra Rai, (2022). Advanced Deep Learning, Packt Publishing, UK.

Bryan Lyon ،Matt Tora, (2023). Exploring Deepfakes: Deploy Powerful AI Techniques for Face Replacement and More with This Comprehensive Guide, New York.

Christian Rathgeb ،Christoph Busch ،Ruben Tolosana ،Ruben Vera-Rodriguez, (2022). Handbook of Digital Face Manipulation and Detection from DeepFakes to Morphing Attacks, Springer International Publishing, New York.

David Sutton (2023). Cyber Security: A Practitioner's Guide, BCS Learning & Development Limited, New York.

De Lima et al. (2020) de Lima O, Franklin S, Basu S, Karwoski B, George A. Deepfake detection using spatiotemporal convolutional networks. Jones & Bartlett Publishers, UK.

Ding et al. (2020) Ding X, Raziei Z, Larson EC, Olinick EV, Krueger P, Hahsler M. Swapped face detection using deep learning and subjective assessment. EURASIP Journal on Information Security, CRC Press, New York.

Fiedelholtz (2021). The Cyber Security Network Guide, Springer Nature, USA.

Floridi (2020) Floridi L. Artificial intelligence, deepfakes and a future of ectypes, Philosophy and Technology, New York.

Frank Millstein, (2023). Convolutional Neural Networks in Python, Paperback, USA.

Goodfellow (2023) Goodfellow I. Generative adversarial nets, Advances in Neural Information Processing Systems, New York.

Guarnera, Giudice & Battiato (2022) Guarnera L, Giudice O, Battiato S. Fighting deepfake by exposing the convolutional traces on images, UK.

Ian Goodfellow, (2021). Deep Learning, MIT Press, USA.

https://www.jpsa.ac.ae/journal/vol15/iss2/2
DOI: 10.69672/3007-3529.1028

32

Ignas Kalpokas ،Julija Kalpokiene, (2022). Deepfakes: A Realistic Assessment of Potentials, Risks, and Policy Regulation, Springer International Publishing, New York.

Joan Daemen & Vincent Rijmen (2022). The Advanced Encryption Standard, Springer Science & Business Media, New York.

Lu Jin (M.S. in information studies) (2023). Cheated by Deepfakes? Deepfake Detection Ability, People's Reactions, and Ethical Implications, University of Texas, New York.

Mariëtte van Huijstee ،Pieter van Boheemen ،Djurre Das ،Linda Nierling ،Jutta Jahnel, Murat Karaboga ،Martin Fatun (2023). Tackling Deepfakes in European Policy, European Parliament, UK.

Maurizio Martellini (2019). Cyber Security: Deterrence and IT Protection for Critical Infrastructures, Springer International Publishing, USA.

Menno van Doorn ،Sander Duivestein ،Thijs Pepping (2022). Real Fake Playing with Reality in the Age of AI, Deepfakes and the Metaverse, Ludibrium Publishers, New York.

Nicholas Kolokotronis & Stavros Shiaeles (2022). Cyber-Security Threats, Actors, and Dynamic Mitigation, CRC Press, USA.

Noah Giansiracusa (2021). How Algorithms Create and Prevent Fake News, Apress, USA.

Peter Trim & Dr Yang-Im Lee (2021). Cyber Security Management: A Governance, Risk and Compliance Framework, Ashgate Publishing, Ltd., UK.

Rajdeep Chakraborty & Anupam Ghosh & Jyotsna Kumar Mandal (2022). Machine Learning Techniques and Analytics for Cloud Security, John Wiley & Sons, USA.

Rebecca J. Blankenship (2021). Deep Fakes, Fake News, and Misinformation in Online Teaching and Learning Technologies, IGI Global, USA.

Sarah Darer Littman, (2022). Deepfake, Scholastic, Incorporated, USA.

Sebastian Raschka, (2022). Machine Learning and Deep Learning with Python, Packt Publishing Ltd, New York.

Xun Yi & Russell Paulet & Elisa Bertino (2021). Homomorphic Encryption and Applications, Springer, USA.

Yuri Diogenes & Dr. Erdal Ozkaya (2019). Cybersecurity – Attack and Defense Strategies, Packt Publishing Ltd, UK.

https://www.jpsa.ac.ae/journal/vol15/iss2/2
DOI: 10.69672/3007-3529.1028