Text k prezentaci na téma deepfake

Tento text je pouhou ukázkou a doporučením interpretace tématu deepfake.

1. snímek - úvod

Tématem dnešní hodiny bude novodobý způsob šíření dezinformací – deepfake. Setkal se již někdo z vás s tímto pojmem/dokázal by ho vysvětlit?

2. snímek – charakteristika

Deepfake je technologie, která za použití umělé inteligence vytváří dojem opravdovosti u upravených videí, obrázků či nahrávek. Nejčastěji se však jedná o kombinaci všech výše zmíněných dohromady. Nově vytvořená videa, případně audia, se poté vkládají do již existujících videí. Algoritmus pracuje s velkým množstvím dostupných dat, například videí, obrázků. Na základě těchto podkladů vytváří výslednou napodobeninu.

Algoritmus, na kterém deepfakes vznikají, se označuje jako deep learning, je to technika strojového učení, kterou jsem vám již takřka popsal. Funguje na principu neuronových sítí, které vidíte na obrázku…Jako další příklad využití deep learning jsou například samořiditelná auta…(učí se značky, objekty, všechny možné situace, které v provozu mohou nastat).

Označení deep = hluboký, je symbolem několika skrytých vrstev, které ve výsledném produktu (deepfake) nejsou vidět. Jedná se o tisíce obrázků, videí, audiozáznamů a dalších…

3. snímek – Shallow fake

Rozdíl oproti shallow fake (mělký, plytký, **povrchní**).

Pojem často zaměňovaný s deepfake. Shallow fake při svém vytváření nevyužívá strojovou techniku učení – deep learning, ale jedná se většinou pouze o triviální úpravy videí jako je zpomalení, zrychlení nebo vystřihnutí části videa z kontextu. Tato úprava je provedena člověkem a není vytvořená umělou inteligencí.

Příkladem shallow fake videa může být americká politička, která vystupovala v televizním pořadu. V následujících dnech se začalo virálně šířit video, ve kterém byla politička zpomalena, a tak vypadala a zněla jako opilá. **Avšak toto není deepfake.**

Obdobně to funguje i u fotografií, které mohou být také jak deepfake tak i shallow fake.

Na obrazcích vidíte koláže od autora Tomáše Břímka, které jsou nedílnou součástí titulních stran časopisu Reflex. Opět se jedná o úpravy fotografií ve photoshopu či jiném příbuzném programu. Všechny vrstvy vidíme a nic nebylo vytvořeno algoritmem, nebylo použité strojové učení počítače – deep learning.

4. snímek – Vznik deepfake

Tvorba deepfake videí v pornoprůmyslu začala v roce 2017. Uživatelé sociální sítě Reddit začali ve velkém množství s tvorbou deepfake pornovideí, ve kterých za pomocí techniky deep learning nahrazovali obličeje aktérů videí známými osobnostmi. Jednalo se o velké množství hereček z Hollywoodu, z populárních seriálů jako Game of Thrones a dalších.

Vypracoval: Radovan Štencel

5. snímek – Výskyt deepfakes

Nejčastější výskyt deepfakes je stále pro zábavu, následován sportem, módou a dalšími odvětvími. Problematický je však užití deepfake technologií v politice. Tento výskyt se začíná navyšovat a může dojít k zneužití této technologie například při volbách.

6. snímek – jak je rozpoznat (kontext)

- Pracovat se zkušenostmi a s kontextem
- Kriticky přemýšlet
- Zaměřit se na emocionální stránku videa
- Zamyslet se nad pravděpodobností daného výroku

7-8. snímek – jak je rozpoznat (technické nedostatky)

- Nepřirozený pohyb očí a nepřirozené mrkání. Je velmi náročné napodobit přirozený pohyb očí tak, aby vypadal dostatečně přirozeně. Taktéž obočí se například může objevit v nečekaných či nepřirozených místech. Moc nízko nebo moc vysoko.
- Nepřirozené výrazy.
- Nepřirozené postavení těla nebo obličeje. Pokud například hlava směřuje jedním směrem, ale nos opačným směrem.
- Vlasy mohou být také nápomocné pro odhalení deepfake videa. Zejména u osob s kudrnatými vlasy nemusí být technologie schopná dostatečně napodobit jejich přirozený pohyb.
- Nedostatek přirozených emocí. Pokud v kontextu k mluvenému slovu nepřicházejí také adekvátní emoce, může se jednat o deepfake.
- Nepřirozené pohyby těla. Pokud se osoba ve videu pohybuje trhaně nebo s nevysvětlitelnými skoky v pohybu, může to být indicie.
- Nepřirozené zbarvení. Může se jednat například o zvláštní odstín pleti na obličeji nebo v okolí očí. Pleť by se také neměla v průběhu videa nijak měnit.
- **DALŠÍ SNÍMEK**
- Nekonzistentní audio. U některých deepfake videí dochází k nedostatečné práci s hlasem a výsledkem může být kovový, nepřirozený až robotický hlas.
- Při zpomalení videa a zaměření se na konkrétní části obličeje, například ústa, lze zkontrolovat, zda se ústa opravdu pohybují adekvátně k mluvenému slovu.
- Pokud se jedná například o deepfake nějaké známé osoby, pravost lze ověřit také vyhledáním podoby dané osoby a následným srovnáním s postavou ve videu.

9. snímek-ukázka jak fungují deepfakes při nedostatku dat

Nejdříve video pusťte jednou, poté doplňte o výklad a opětovně pusťte/pozastavte.

Toto video nám demonstruje, co se stane, když technika strojového učení (deep learning) nemá dostatek podkladů pro tvorbu samotného deepfake videa. Pokud se žena otočí, tak se její obličej změní do původního stavu. Pokud se však dívá přímo na nás, její obličej je upraven. Technologie tedy pravděpodobně neměla fotky ženy z boku.

10. snímek-Deepfake video, prezident Zeman

Video slouží již pouze pro ukázku samotného deepfake videa.

Vypracoval: Radovan Štencel