

SVEUČILIŠTE U ZAGREBU
FAKULTET ELEKTROTEHNIKE I RAČUNARSTVA

SEMINAR

**Detekcija uvjerljivog krivotvorenog
sadržaja na društvenim mrežama
uporabom GAN-ova**

Barbara Kos, Matija Pavlović

Voditelj: *prof. dr. sc. Tomislav Hrkać*

Zagreb, prosinac 2023.

**Detekcija uvjerljivog krivotvorenog sadržaja na društvenim mrežama
uporabom GAN-ova**

Sažetak

Unesite sažetak na hrvatskom.

Ključne riječi:

Abstract

Abstract

Keywords:

SADRŽAJ

1. Uvod	1
1.1. Pojava	1
1.2. Opasnosti	1
2. Razrada	2
2.1. Metode stvaranja	2
2.2. Artefakti	2
2.3. GAN-ovi	2
2.3.1. Generator	2
2.3.2. Diskriminator	2
2.4. Primjena GAN-ova za detekciju deepfake-ova	2
3. Rezultati i rasprava	3
4. Zaključak	4
5. Privitci	5

1. Uvod

1.1. Pojava

Prva poznata pojava pojma "deepfake" datira iz prosinca 2017. godine kada je korisnik Reddita osnovao "subreddit" pod nazivom "r/deepfakes". Ovaj podforum uglavnom je sadržavao pornografski sadržaj u kojem su izmjenjena lica kako bi nalikovali poznatim osobama. Ovaj fenomen predstavlja tehnološki napredak, ali istovremeno izaziva zabrinutost zbog potencijalne zloupotrebe.

Takvi sadržaji često su prikazivali poznate i utjecajne osobe u situacijama koje se nikada nisu dogodile. Neki od poznatijih primjera uključuju papu Franju, bivšeg predsjednika SAD-a Donalda Trumpa te druge javne ličnosti. Ovaj trend je brzo stekao popularnost, često zbog senzacionalizma i šoka koji izaziva.

Unatoč negativnom kontekstu u kojem se često spominje stvaranje deepfake sadržaja, važno je napomenuti da postoji i pozitivan aspekt primjene ove tehnologije. Naime, deepfake tehnologija može se koristiti u edukativne svrhe, kao što je stvaranje videa u kojima poznate osobe, poput Davida Beckhama, podižu svijest o globalnim problemima poput malarije na različitim jezicima.

Osim toga, deepfake tehnologija nalazi primjenu u umjetnosti i zabavi, npr. u stvaranju scena u filmovima nakon smrti glumaca ili njihova digitalnog starenja. Također, postoji značajan broj deepfakeova čija je svrha isključivo humoristična, a takvi se sadržaji često viralno šire društvenim mrežama.

Važno je razumjeti da deepfake tehnologija nosi sa sobom i etičke izazove te da njezina primjena zahtijeva odgovornost kako bi se izbjegla potencijalna šteta i zloupotreba.

1.2. Opasnosti

2. Razrada

2.1. Metode stvaranja

2.2. Artefakti

2.3. GAN-ovi

GAN-ovi su predstavljeni 2014. godine od strane Ian Goodfellowa i njegovih kolega. Osnovna ideja iza GAN sustava jest postojanje dva glavna dijela mreže: generator i diskriminator. Tijekom treninga, generator i diskriminator se natječu jedan protiv drugoga. Generator pokušava poboljšati svoje sposobnosti generiranja tako da vara diskriminator, dok diskriminator nastoji postati sve bolji u razlikovanju pravih podataka od lažnih. Ovaj suparnički proces dovodi do poboljšanja kvalitete generiranih podataka tijekom vremena.

2.3.1. Generator

Generator GAN-a je ključni dio sustava čija je zadaća stvaranje novih uzoraka ili podataka koji bi trebali biti što je moguće sličniji stvarnim primjerima iz skupa podataka na kojem je mreža trenirana. Na ulaz generatora dovodi se nasumični šum, a zatim se on izmjenjuje u izlaz koji nalikuje podacima iz skupa za treniranje. Uvođenjem nasumičnog šuma, postizemo generiranje raznolikih podataka, uzorkovanjem iz različitih točaka ciljne distribucije. Generator dakle kreira sadržaj suparničkim pristupom, pokušava prevariti diskriminator.

2.3.2. Diskriminator

2.4. Primjena GAN-ova za detekciju deepfake-ova

3. Rezultati i rasprava

Rekreirali smo taj i taj rad, ostvarili te i te rezultate, možemo ih koristiti za to i to

4. Zaključak

navesi budući rad, detekcija u videima itd.

5. Privitci