Univerza v Ljubljani
Fakulteta za računalništvo in informatiko

Slikovna forenzika

dr. Borut Batagelj
sodni izvedenec

FORENZIČNO-KRIMINALISTIČNO TEHNIČNE PREISKAVE
preiskave fotografij ter posnetkov video kamer (video nadzor)

član Stalnega strokovnega telesa za forenziko
pri Strokovni svet za sodno izvedenstvo, sodno cenilstvo in sodno tolmačenje,

Laboratorij za računalniški vid, UL-FRI

raziskovalno področje: slikovna biometrija

30. marec
2023

Zakon o sodnih izvedencih, sodnih cenilcih in sodnih tolmačih (ZSICT)

2. člen (status)

(1) Sodni izvedenci so osebe, imenovane za neomejen čas s pravico in dolžnostjo, da sodišču na njegovo zahtevo podajo izvid in mnenje glede **strokovnih vprašanj**, za katera tako določa zakon ali glede katerih sodišče oceni, da mu je pri njihovi presoji potrebna pomoč strokovnjaka.

Strokovna vprašanja

IZBOLJŠAVA

Analiza/ izboljšava posnetka

- gibanje sumljive osebe
- dogajanje na posnetku

PRISTNOST

Ali so posnetki/fotografije pristne?

- določiti, potrditi snamalno napravo
- določiti čas nastanka

IDENTIFIKACIJA

Ali se na posnetku nahaja določena oseba?

- obrazne karakteristike
- poškodbe, znamenja
- višina
- dolžina obuvala

Ali se na posnetkih pojavljajo iste osebe?

Pristnost posnetkov

- Obdani z vizualnimi podobami
- Zaupanje: nekoč in danes
  - Rumeni tisk, modna industrija, mediji
  - Znanstvene revije
  - Politična propaganda
  - Ponarejanje dokazov na sodišču
  - Različne slikovne potegavščine (email)

- Ponarejanje je dandanes enostavno
  - Razširjenost digitalnih kamer
  - Dostopnost do programske opreme za manipulacijo
- Naloga digitalne forenzike: Povrniti zaupanje

Globoki ponaredki

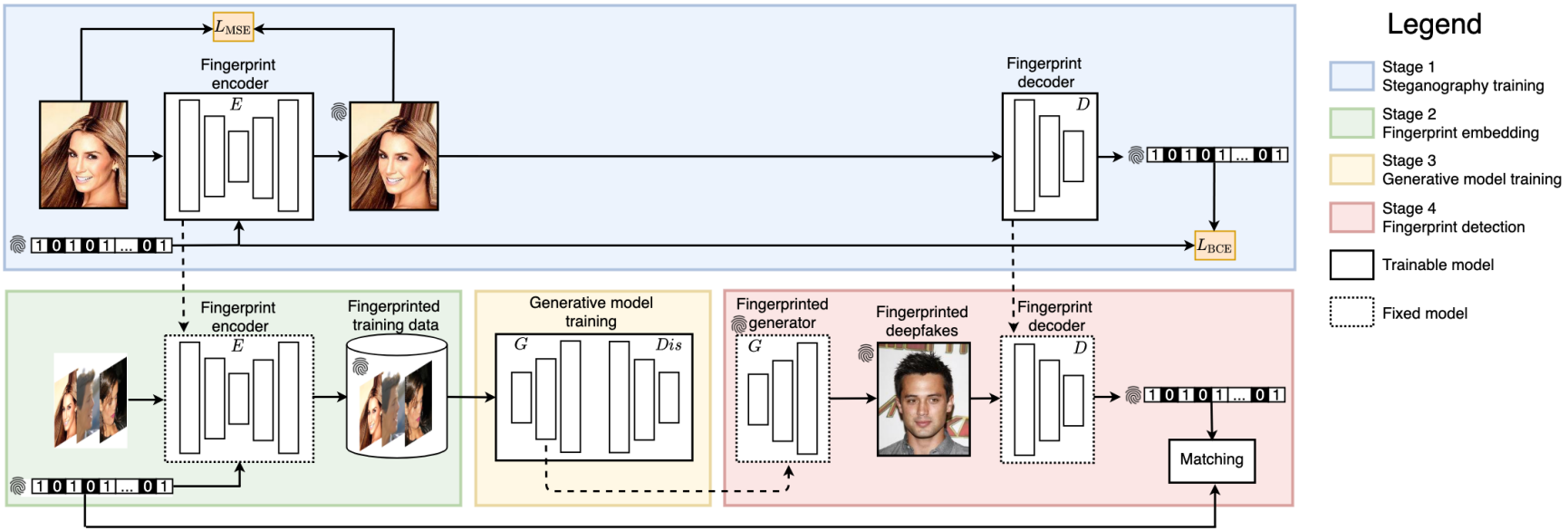
Vir: https://www.youtube.com/watch?v=cQ54GDm1eL0

Metode za odkrivanje ponaredkov

- Aktivne metode
  - Vodni žig
  - Digitalni podpis
  - CAI in C2PA

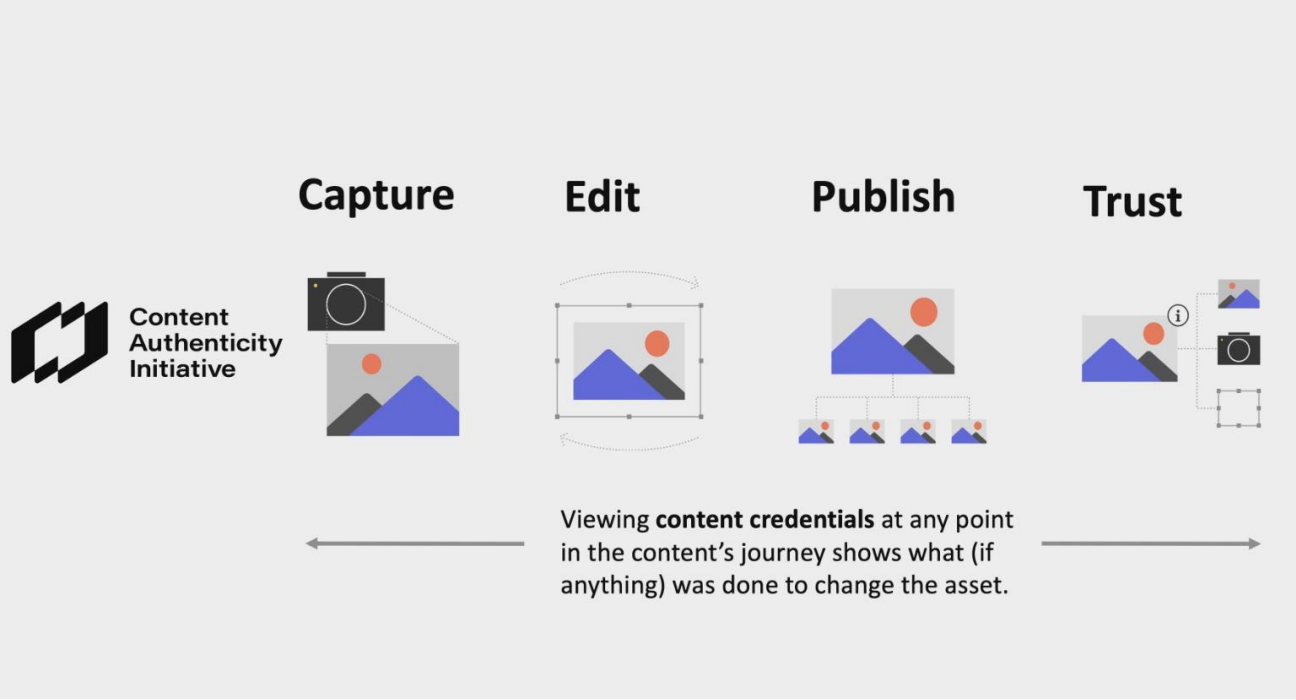
- Pasivne metode
  - Ni vidnih sprememb
  - Spremeni se notranja statistika podatkov
    - Klasične metode
    - Metode na osnovi globokega učenje

Aktivne metode: Digitalni žig



Yu et al. 2021, Artificial Fingerprinting for Generative Models: Rooting Deepfake Attribution in Training Data

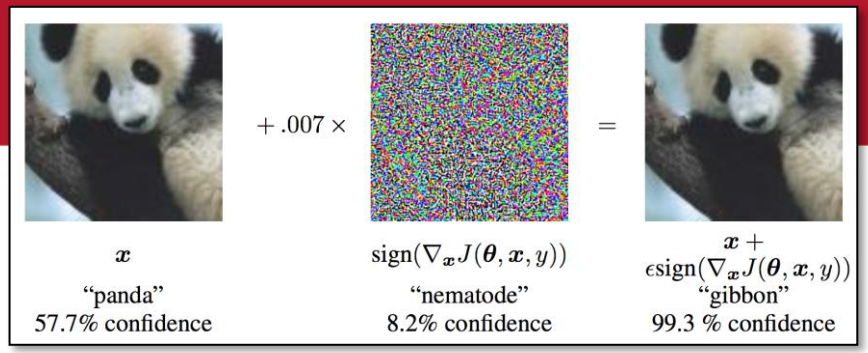
Adobe: Pobuda za pristnost vsebine



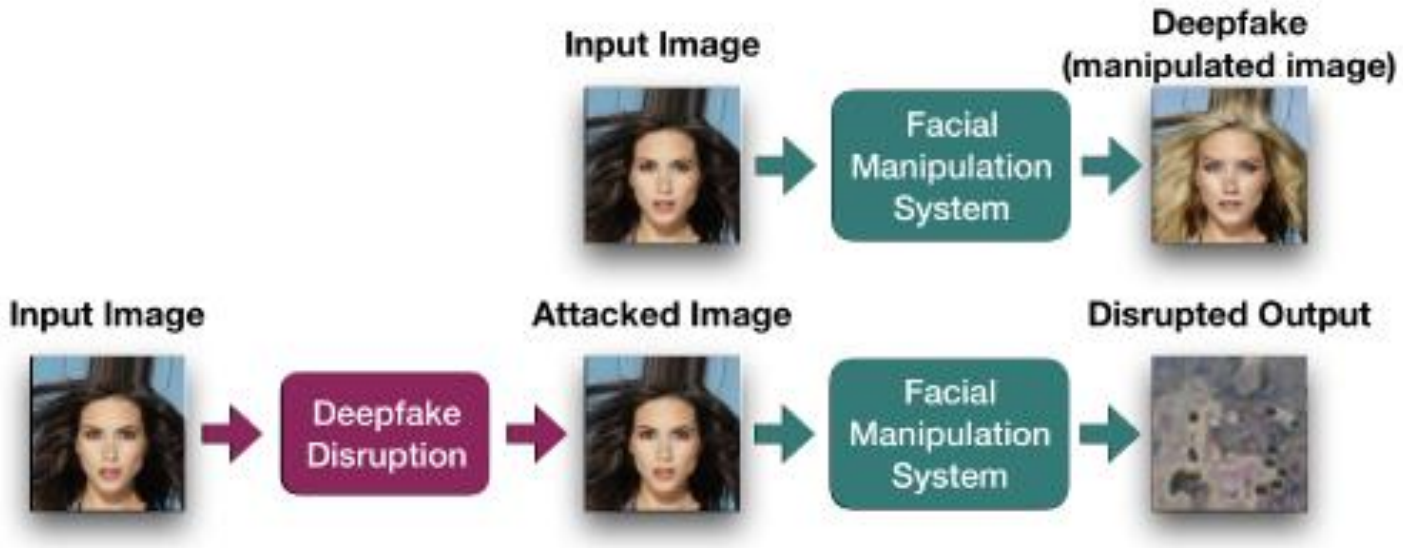
The Coalition for Content Provenance and Authenticity (C2PA)

https://c2pa.org/

Aktivne metode: Nasprotniški napad



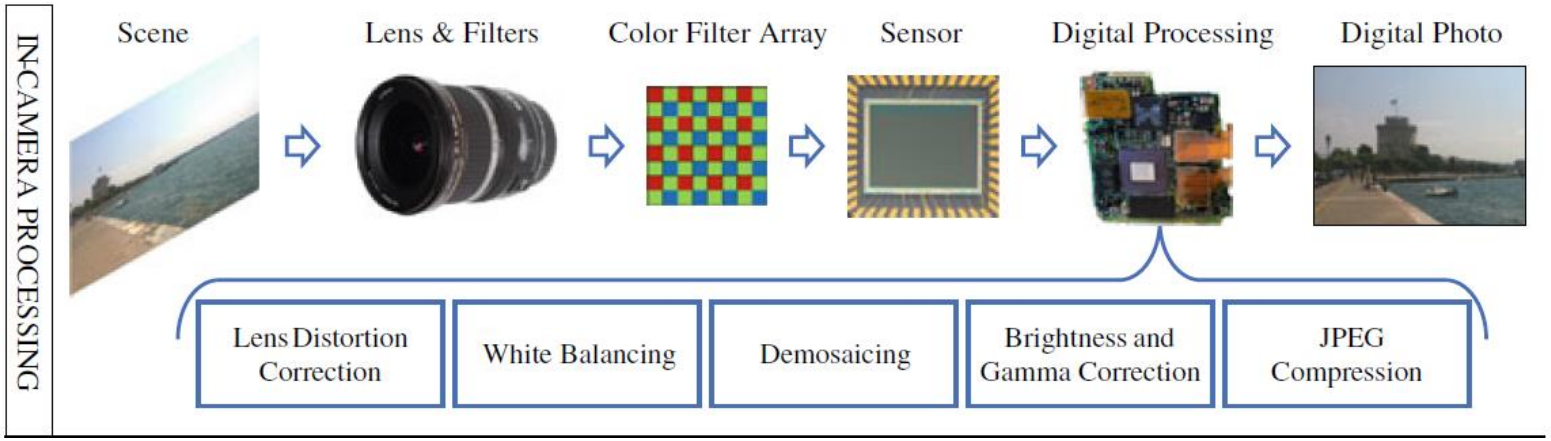
Goodfellow et al. 2015



Ruiz et al. 2020, Disrupting Deepfakes: Adversarial Attacks Against Conditional Image Translation Networks and Facial Manipulation Systems

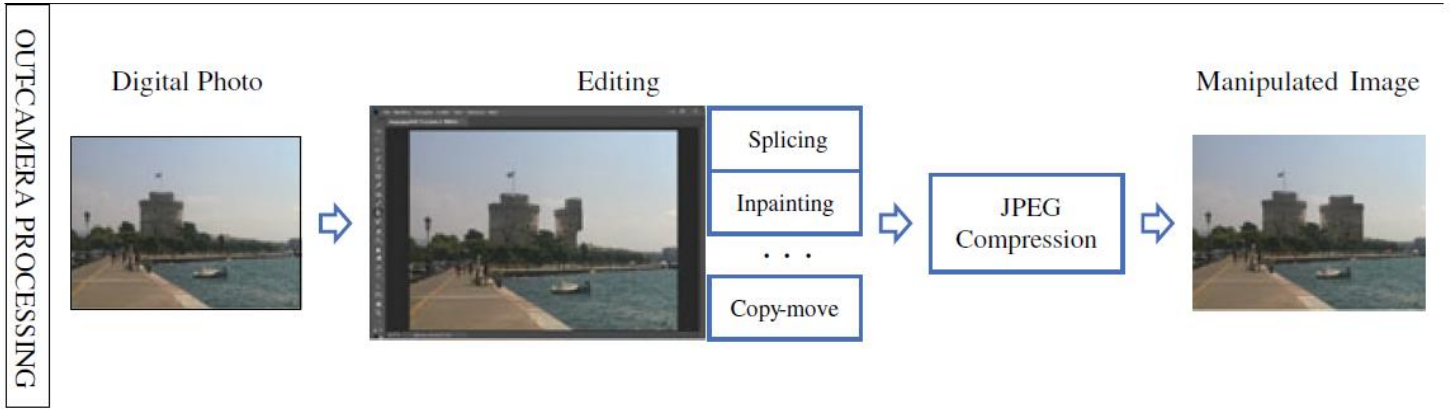
V kameri

- Sliko zajeme napava (značilnosti)
- Vsako procesiranje spremeni prvotni vzorec



Izven kamere

- Slika je lahko spremenjena na različne načine: dodajanje/brisanje ali podvajanje elementov
- Na koncu se slika običajno ponovno shrani v JPG format



Pasivne metode

1. Na nivoju slikovnega elementa
2. Na nivoju formata (izgubno stiskanje)
3. Lastnosti kamere: leče, senzor, obdelava na čipu
4. Fizikalne lastnosti (objekt-osvetlitev-kamera)
5. Geometrijske lastnosti (fotogrametrija)

Hany Farid, Image Forgery Detection, A survey

Na nivoju slikovnega elementa

- Druge analize
  - identifikacija : DNK, prstni odtis, obraz
  - odontologija : zobovje
  - entomologija = nauk o žuželkah : insekti
  - geologija : prst, zemlja

- Digitalna računalniška forenzika: bit

- Forenzična analiza slik: **slikovni element**

Na nivoju slikovnega elementa
Kloniranje

- Ponovitev dela slike
  - Pretiravanje, zavajanje
  - Prekrivanje osebe ali objekta
- Računska kompleksnost
  - rešitev: DCT ali PCA s podobnimi koeficienti

Objavljena slika

Originalna slika

Na nivoju slikovnega elementa

Ponovno vzorčenje

- Nova kompozicija
  - Del slike povečamo/zmanjšamo, zasukamo
    - Ponovno vzorčenje => korelacija sosednih elementov

Na nivoju slikovnega elementa

Ponovno vzorčenje

- Nova kompozicija
  - Del slike povečamo/zmanjšamo, zasukamo
    - Ponovno vzorčenje => korelacija sosednih elementov

Na nivoju slikovnega elementa
Spajanje

- Dve ali več slik v eno
- Spremembe na robu vizualno niso opazne
- Spremenijo pa se Fourierove statistike višjega reda

Oprah

v telesu Ann Margret

Na nivoju slikovnega elementa

Statistika

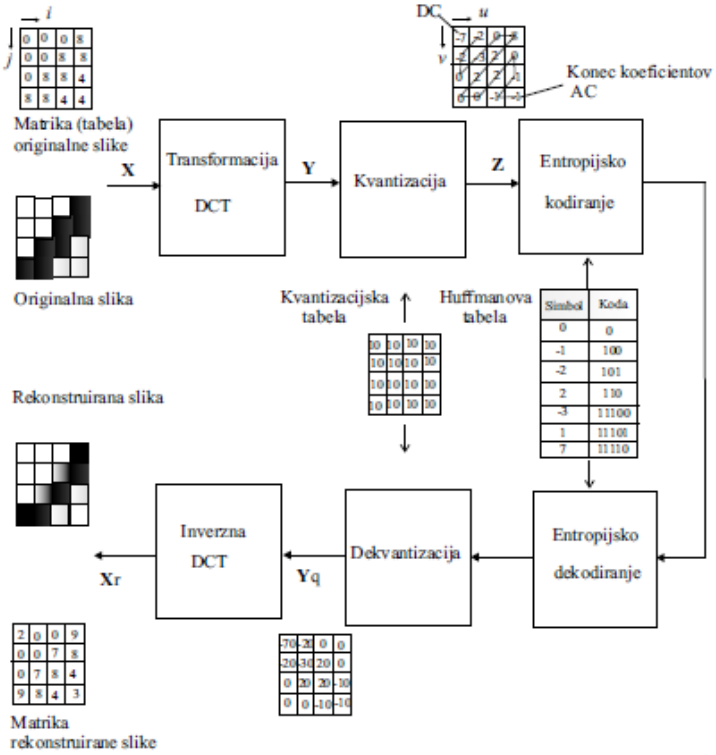
- Veliko kombinacij slik: 256^{n*n}, n=10 ≈ 10^{240}
- Zelo malo takšnih, ki imajo nek pomen
- Fotografije sledijo določenim statistikam
  - Izračunane statistike => ali je bila slika spremenjena
    - Lahko se ugotovi
      - povečavo, filtriranje
      - pravo fotografijo ali rač. generirano
      - skrita sporočila (steganografija)
        - » co-occurrence matrike (matrika sočasne pojavitve)

Na osnovi formata

- Prvo pravilo forenzične analize
  - Ohraniti dokaze
- Ali je izgubni format JPEG težava?

- Lastnosti JPEG se lahko uporabi za forenzično analizo:
  - Kvantizacija
  - Dvojna kompresija
  - Nivojska analiza napake
  - JPEG bloki

Na osnovi formata
JPEG Kvantizacija



- Shema stiskanja na nivoju fotoaparata
- Določimo izvorno napravo
- Razlike znotraj ene kamere (nastavitve)
- Prekrivanje med različnimi kamerami

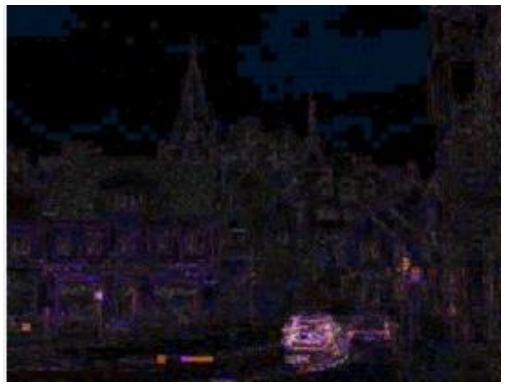
= digitalna slikovna balistika

➡ izvor slike lahko potrdimo ali ovržemo

Na osnovi formata
Dvojni JPEG

- Manipulacija -> ponovno shranjevanje
  - original v JPG
  - sprememba v JPG
- Dvojna kompresija dodanih delov
  - Nepravilnosti pri izgubnem stiskanju
  - Služi za dokaz manipulacije

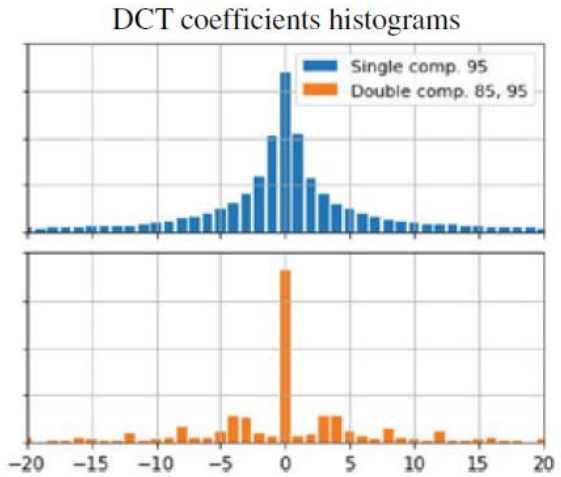
- Na celotni sliki ne more dokazati zlonamerne spremembe
  - sliko smo lahko samo ponovno shranili

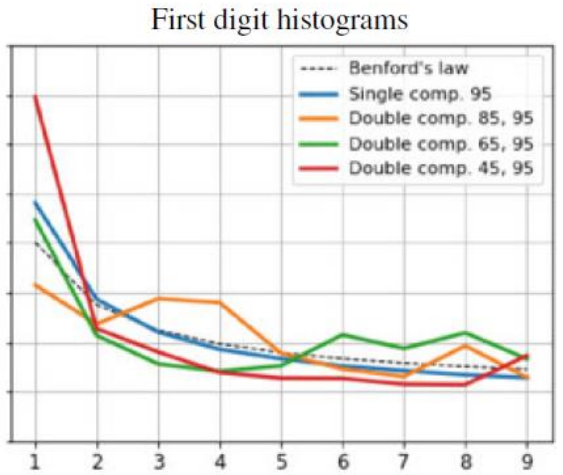


Dvojna kompresija na delu slike!

Kompresija

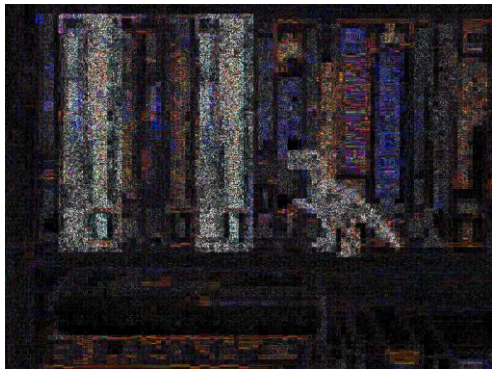
- Dvojna kompresija
  - Različni artefakti
  - koeficienti DCT so v skladu z Benfordovim zakonom
  - Nivojska analiza napake (ELA)
    - Išče se tako imenovane anomalije, ki se pojavijo, če kompresiramo sliko z enako kompresijo dvakrat (JPEG ghost)
    - Izvede se tako, da se zakompresira sliko z različnimi faktorji in se pogleda razlike





Na osnovi formata
Nivojska analiza napake

- Zaznamo dele na sliki, ki imajo različno kompresijo
  - Podobne površine imajo isto stopnjo kvalitete
  - Nov del ima drugačno stopnjo kvalitete
  - Različne dele vidimo predstavljene z nivoji svetlosti
  - Predmeti ki so bili dodani zadnji so svetlejši

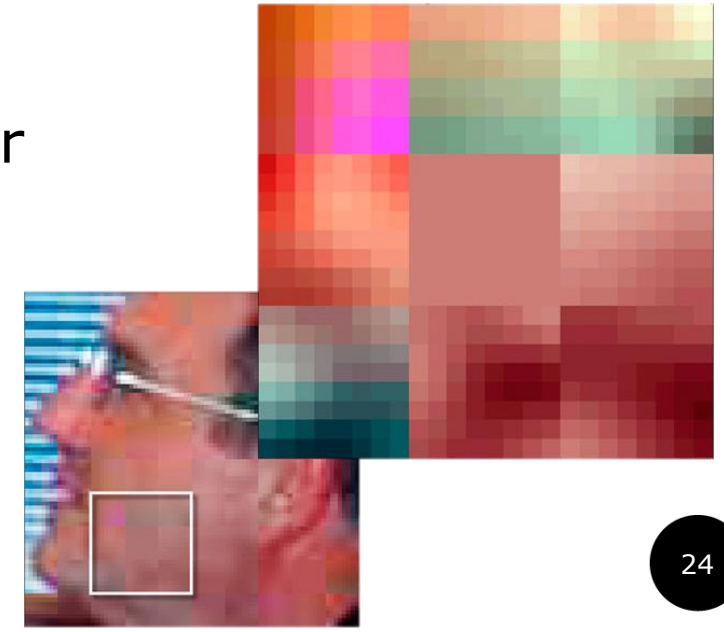


Error level analysis - ELA

Na osnovi formata
JPEG bloki

- Osnova JPEG kompresije so DCT bloki
- Obdelava (transformacija+kvantizacija) se vrši na blokih velikosti 8×8
- Na robovih blokov nastane popačenje
- Če sliko spremenimo se to pozna na robnih točkah

- Izračuna se lastnosti iz delov kjer ni popačenja
  => določitev spremenjenih regij



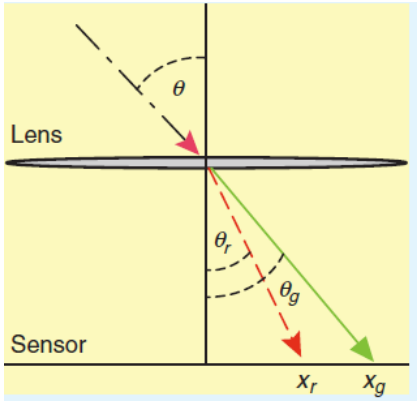
Na osnovi kamere

- Utori na izstrelku
  - povezava z orožjem

- Lastnosti kamere se odražajo na slikah
  - Kromatična aberacija
  - Barvna matrika
  - Odziv kamere
  - Šum senzorja

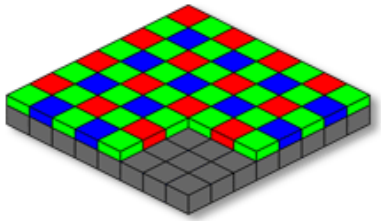
Na osnovi kamere

Kromatična aberacija

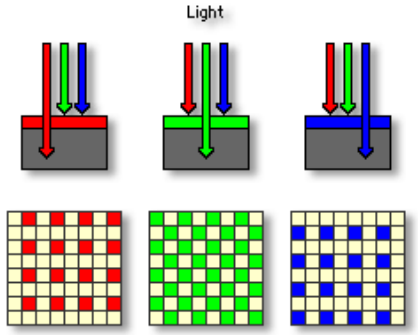


- Lom svetlobe različnih valovnih dolžin
- Primerjava barvnih kanalov
  - Vektor odmikov R in G kanala

Na osnovi kamere
Barvna matrika

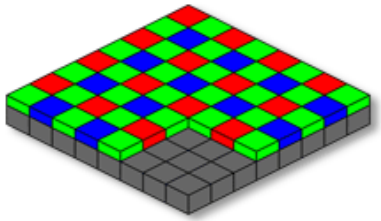


Color Filter Array Sensor

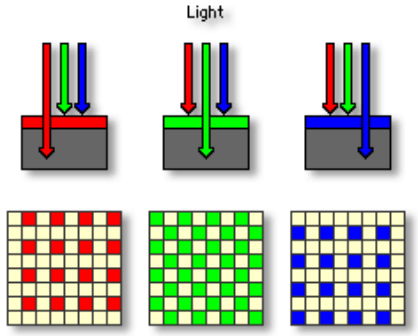


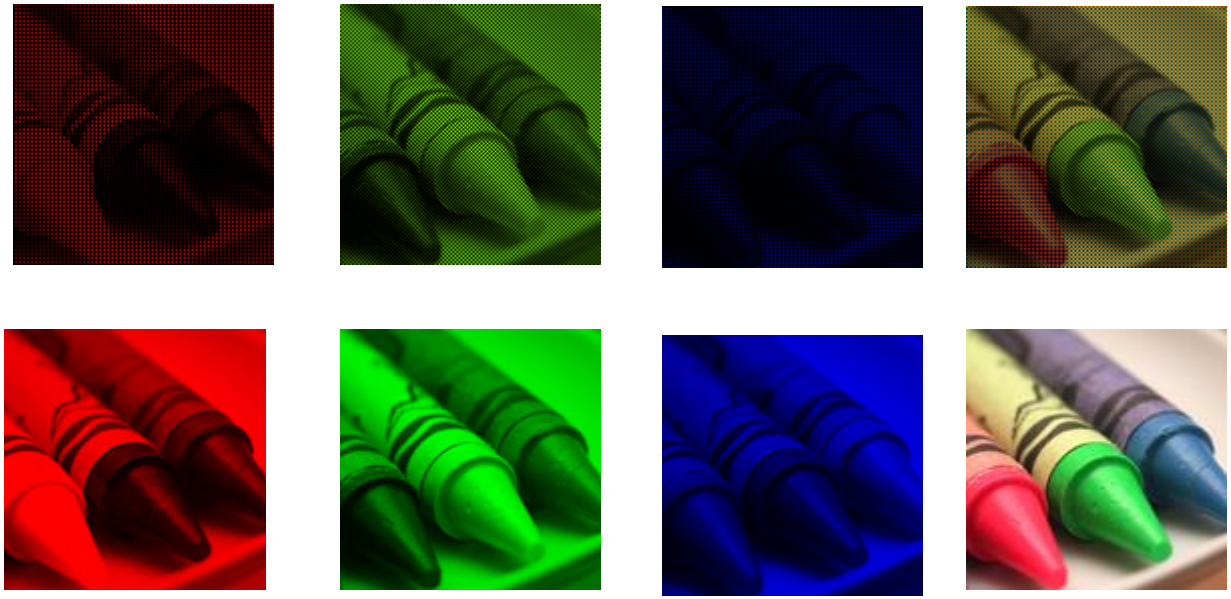
- Barvna slika: 3 barvni kanali RGB
- En svetlobni senzor: različne barve s pomočjo barvnega filtra (CFA)
- Vsak slikovni element senzorja svoj filter
  - Za vsak slik. element ena barva
  - Ostali dve se izračunata s pomočjo interpolacije (demozaičenje)
- Interpolacija pusti na sliki sled
  - Statistična odvisnost v vsakem barvnem kanalu
  - Filter se ponavlja -> vzorec je ponavljajoč
  - Odstopanje od vzorca ➡ sprememba

Na osnovi kamere
Barvna matrika



Color Filter Array Sensor

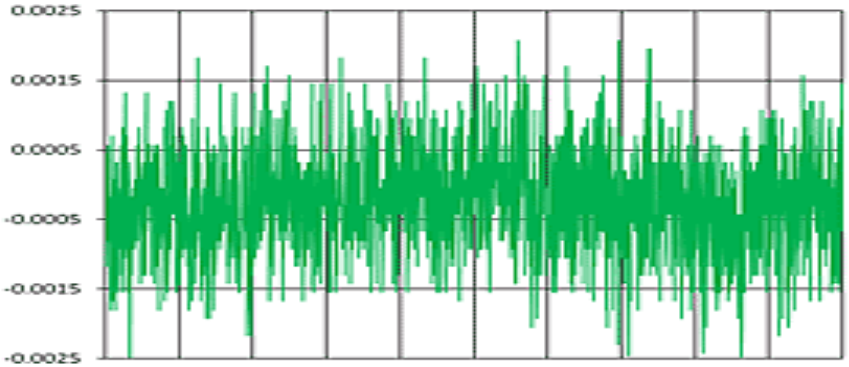




Na osnovi kamere
Odziv kamere

- Senzorji so večinoma linearni
  - Odvisnost med količino svetlobe in vrednost pripadajočega slik. elementa = linearna
- Nelinearni filter za izboljšanje slike
  - Vseeno se ohrani odvisnost
- Izračun odzivne funkcije
  - Preiskovanje po sliki
  - Odziv odstopa ➡ spremembe

Na osnovi kamere
Šum senzorja

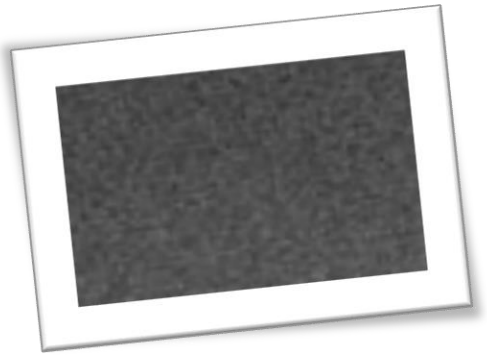


- Slika na sezorju
- Obdelave:
  - Kvantizacija
  - Belina
  - Interpolacija
  - Barvna korekcija
  - Gama korekcija
  - Filtriranje
  - JPEG stiskanje

Vse to pusti določeno sled
- je bila slika obdelana
- iz katere naprave

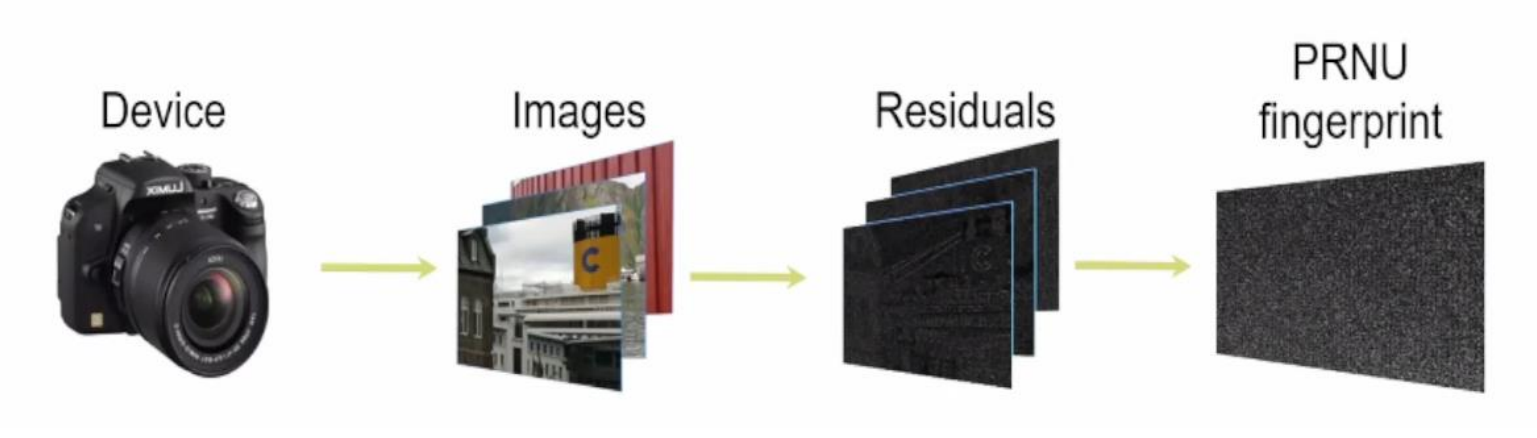
PRNU kot prstni odtis kamere

- Nastane zaradi majhnih odstopanj pri izdelavi senzorjev
- Lahko ga modeliramo kot dodan šum
- Je prisoten v vsaki sliki in konstanten skozi čas
- Določa posamezno napravo
- Močno orodje za odkrivanje izvora



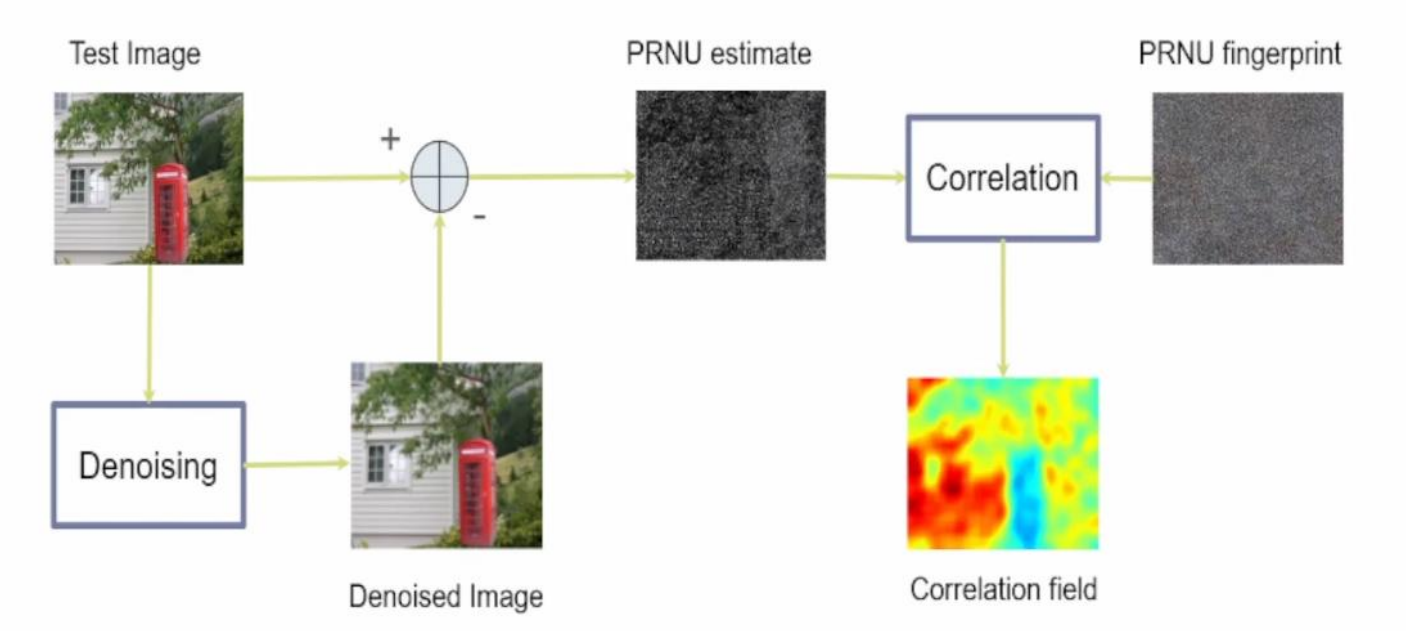
Detekcija na podlagi šuma

- Vzorec šuma (PRNU) lahko izračunamo (natančno določimo) s pomočjo 150-200 slik iz naprave



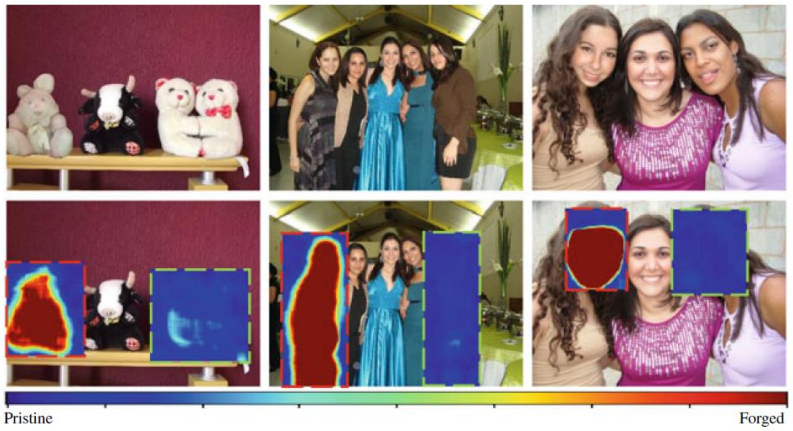
Detekcija na podlagi šuma

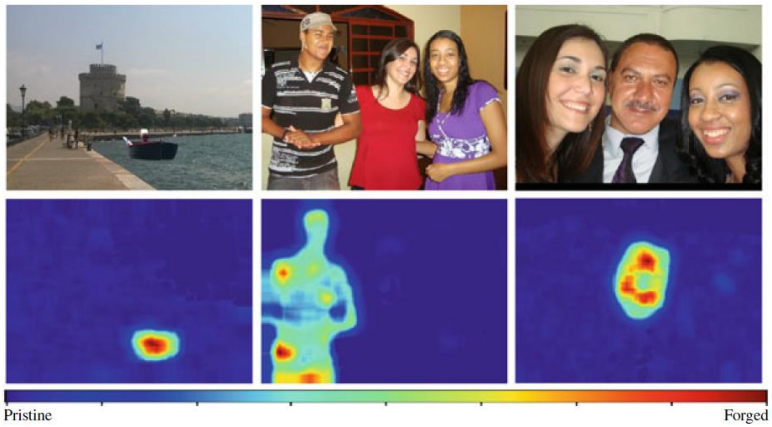
- Ocena šuma iz ene slike



Metoda na osnovi šuma kamere

- Poznamo dva pristopa:
  - Sami določimo regijo (izračun značilnega vzorca se računa na drugih regijah)
  - Izračun se izvrši na celotni sliki





Zakoni fizike

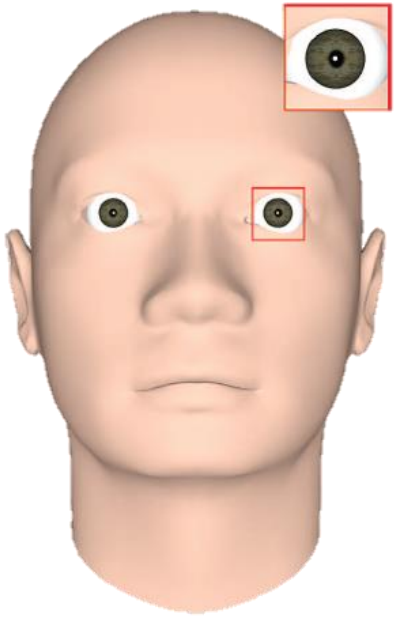
- Dve slavni osebi se sprehajata po plaži
- Lahko, da se osebi sploh ne poznata
  - Izrezane in dodane v okolje
- Težko je zagotoviti svetlobnim pogojem
  - Enake za obe osebi
  - Skladne z okoljem

April, 2005

- Tehnike
  - Usmerjenost svetlobe (2D, 3D)
  - Osvetlitev prostora

Fizika

Usmerjenost svetlobe (3D)



– Pomagamo si z odbojem v očeh
  • Primerjamo za različne ljudi

Fotogrametrija

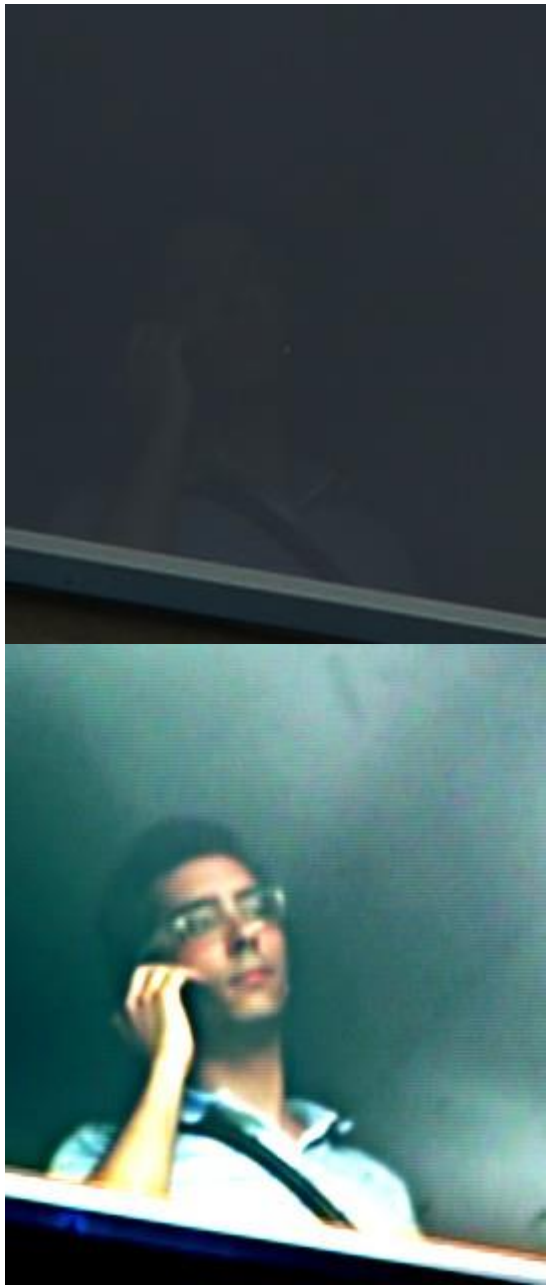
- Transformacija slike
  - Meritve


  - Razberemo oznako/besedilo


Ocena višine (SVM)

Druge tehnike

1. Posvetlitev (nočni posnetki)
2. Povečava kontrasta
3. Odstranitev šuma
4. Stabilizacija videa
5. Popravek premika
6. Razvijanje 360 posnetka
7. Korekcija napak leč



izitru https://www.izitru.com

AMPED SOFTWARE http://ampedsoftware.com/

Manipulacija - video

Vir: https://www.youtube.com/watch?v=iyiOVUbsPcM

Porast globokih ponaredkov (DeepFakes)

Ideja ni nova.

1997 2014 2015 2016 2017 2018 2019 2020 2021

Video rewrite: Driving visual speech with audio" SIGGRAPH (1997)

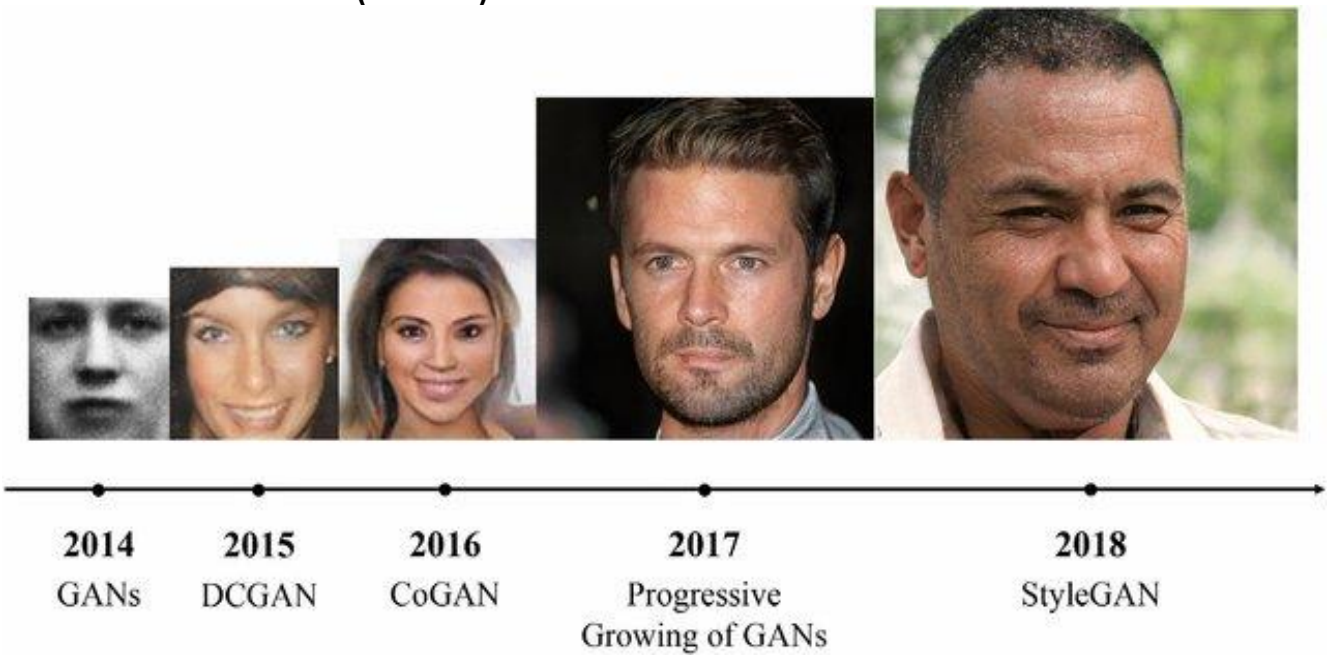
Nvidia, "Unsupervised image-to-image translation", NIPS (2017)

November 2017: Reddit user named **deepfakes** started sharing face-swapping pornographic videos

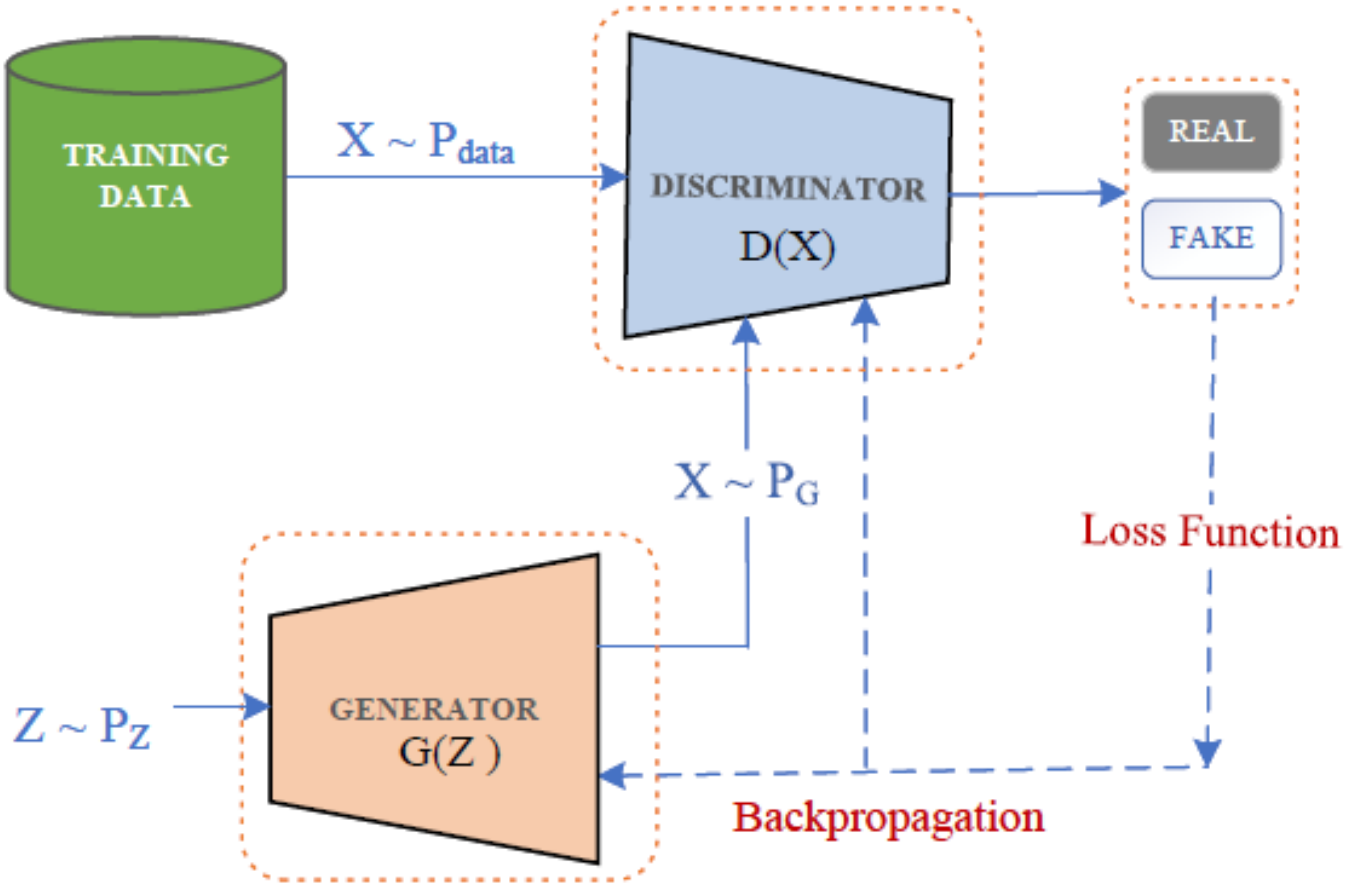
Porast globokih ponaredkov (DeepFakes)

1997 2014 2015 2016 2017 2018 2019 2020 2021

Goodfellow et. al.,
"Generative ***Adversarial*** Networks
" (GANs)



GAN


TRAINING DATA
X ~ P data
DISCRIMINATOR
D(X)
REAL
FAKE
X ~ P G
Loss Function
Z ~ P Z
GENERATOR
G(Z)
Backpropagation

Hitro širjenje

4 pomembi faktorji:

1. Socialna omrežja/platforme

   Facebook, instagram, twitter, youtube, snapchat, LinkedIn, pinterest, tiktok, WeChat

2. Procesorska moč : GPU

   Omogoča poganjanje zapletenih algoritmov na katerih slonijo AI metode

3. Tehnologija globokega učenja

   Učenje s pomočjo globokih konvolucijskih mrež

4. Prosto dostopna programska oprema na odprtih platformah: GitHub

   FaceSwap, FAKEAPP, ZAO

Metode za manipulacijo obraza

- Generiranje novega obraza
- Zamenjava identitete
- Manipulacija določenih atributov
- Zamenjava izraza (ustnice)
- Druge manipulacije

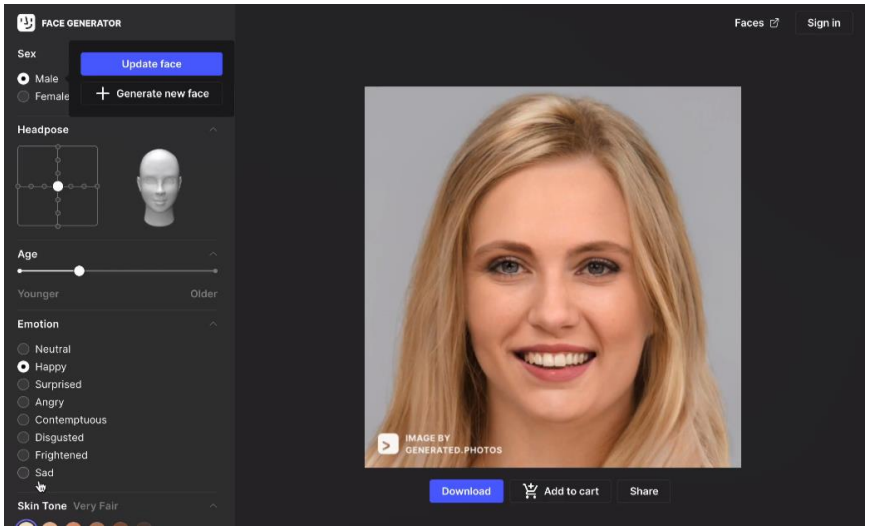
Generiranje novega obraza

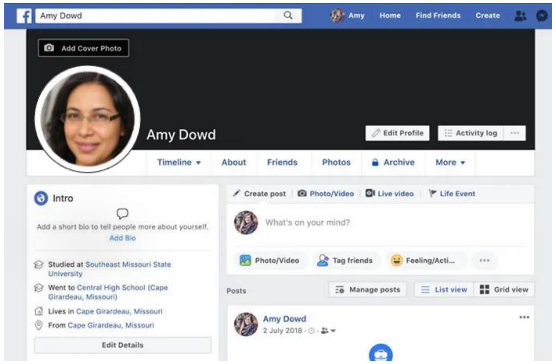


- [This Person Does Not Exist - Free AI Face Generator]
- https://generated.photos/face-generator

Uporabnost:
- 3D modeliranje (igre)
- Filmska industrija

Nevarnosti:
- lažni profil na soc. omrežjih
  - Letno odkrijejo več tisoč lažnih računov.





Zamanjava identitete

- S pomočjo rač. grafike
- S pomočjo globokih nevronskih modelov

Izvorna slika: identiteta ponorna slika: atributi Generirana slika

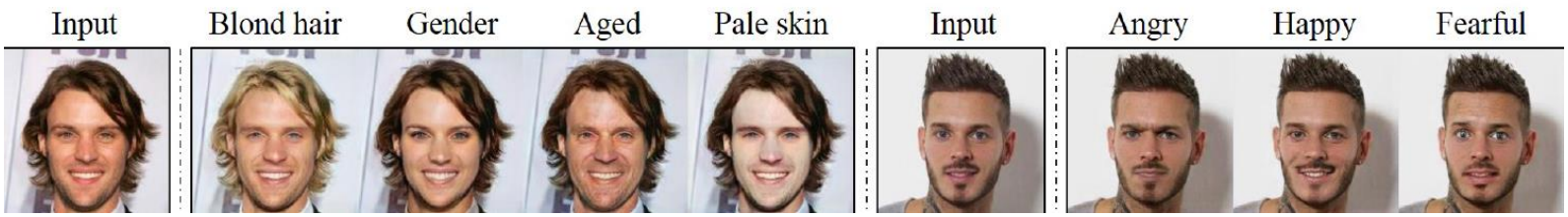
Uporabnost:
- Zabavna industrija
- Filmska industrija

Nevarnosti:
- Lažni pornografski posnetki
- Goljufije

Y. Zhu, Q. Li, J.Wang, C.-Z. Xu, and Z. Sun, “One shot face swapping on megapixels,” in Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2021, pp. 4834–4844.

Manipulacija določenih atributov

- Sprememba frizure, barva kože, spola, starosti

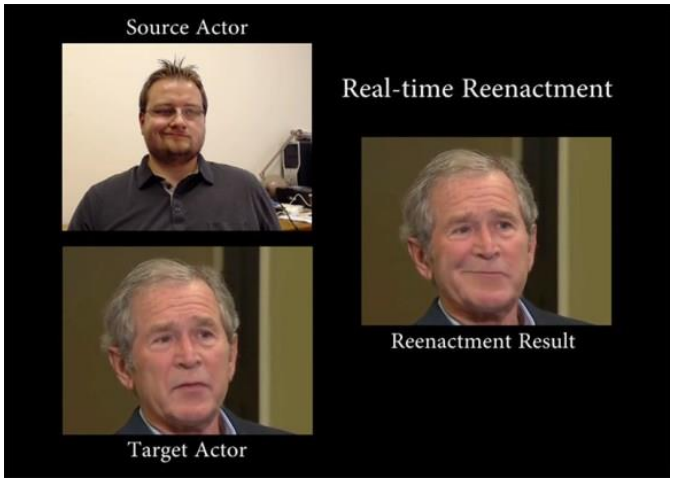


Uporabnost:
- Virtualno okolje
  - Preizkus modnih dodatkov
  - Ličila
  - Kozmetika
  - Pričeska

Mobilna aplikacija:
FaceApp

Zamenjava izraza (ustnice)

- Rekonstrukcija obraza
- Tehnike:
  - Face2Face
  - Neural Textures

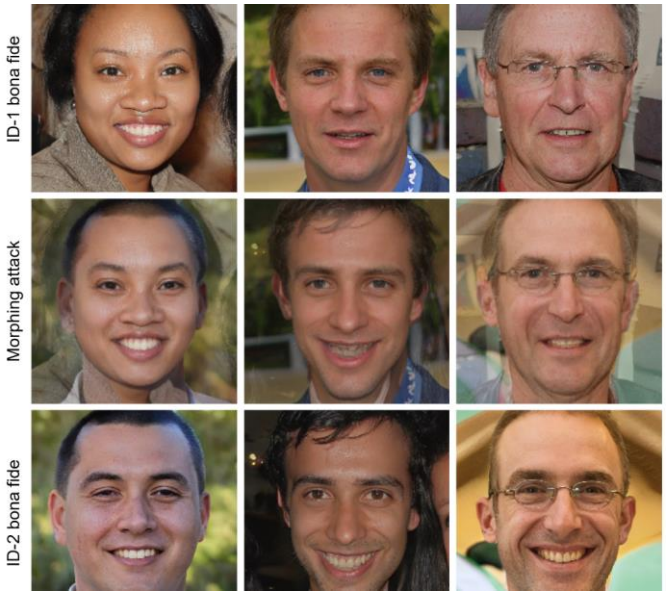


https://www.youtube.com/watch?v=ohmajJTcpNk

Goljufija
- Nekdo reče kar ni rekel

Druge manipulacije

- Preoblikovanje obraza (morphing)
  - Ustvarjanje umetnih biometričnih vzorcev obraza
- Anonimizacija (deidentifikacija) obraza
  - Odstranitev identitete (ali določne značilnosti) iz obraza

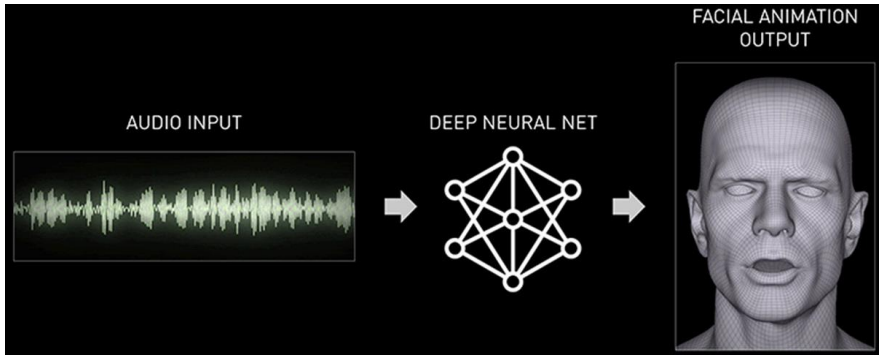


Zamenjava izraza

Iz zvoka v video

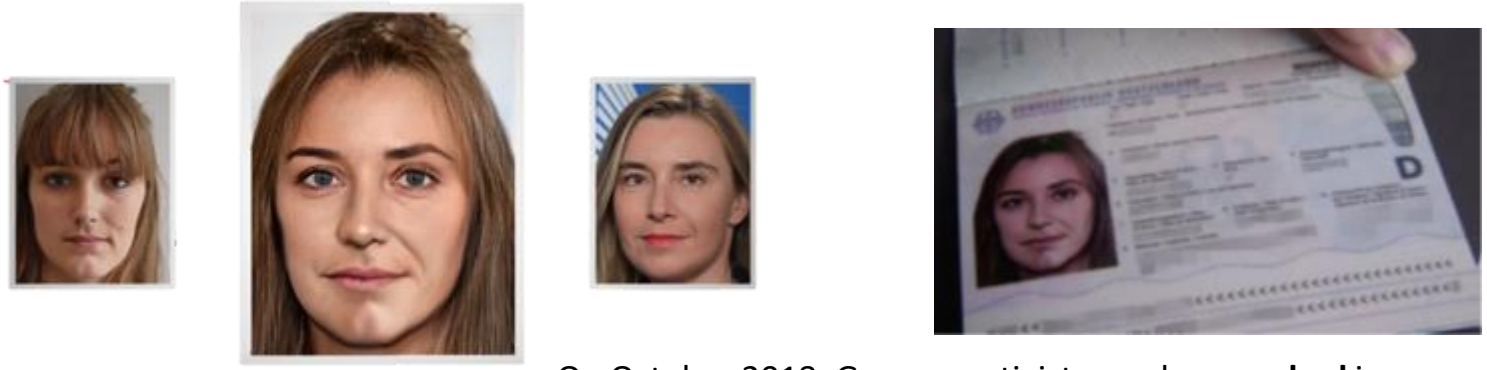
Iz besedila v video

Sinhronizacija ustnic na podlagi novega besedila ali govora



Face Morphing

- The magic passport
- Report by the Slovenian Police [Torkar 2021]
  - in the last 12 months more than 40 morphing cases were detected at Airport Police in Ljubljana (2021)



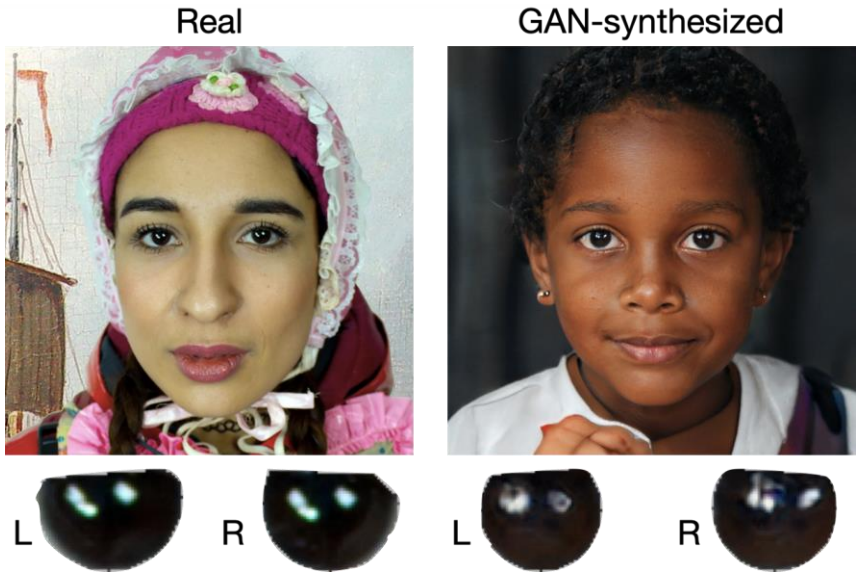
On October 2018, German activists used a **morphed** image of **Federica Mogherini** (High Representative of the European Union for Foreign Affairs and Security Policy) and a member of their group to get a genuine German passport.

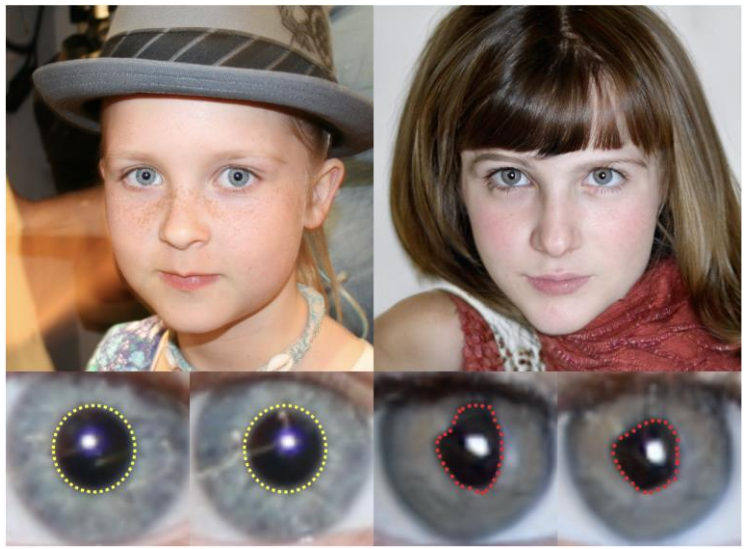
Detekcija globokih ponaredkov

- Vidne napake na generiranih oz. spremenjenih slikah
  - Artefakti na sliki (ozadje)
  - Zobje pri nefrontalnih obrazih

Detekcija globokih ponaredkov (slike)

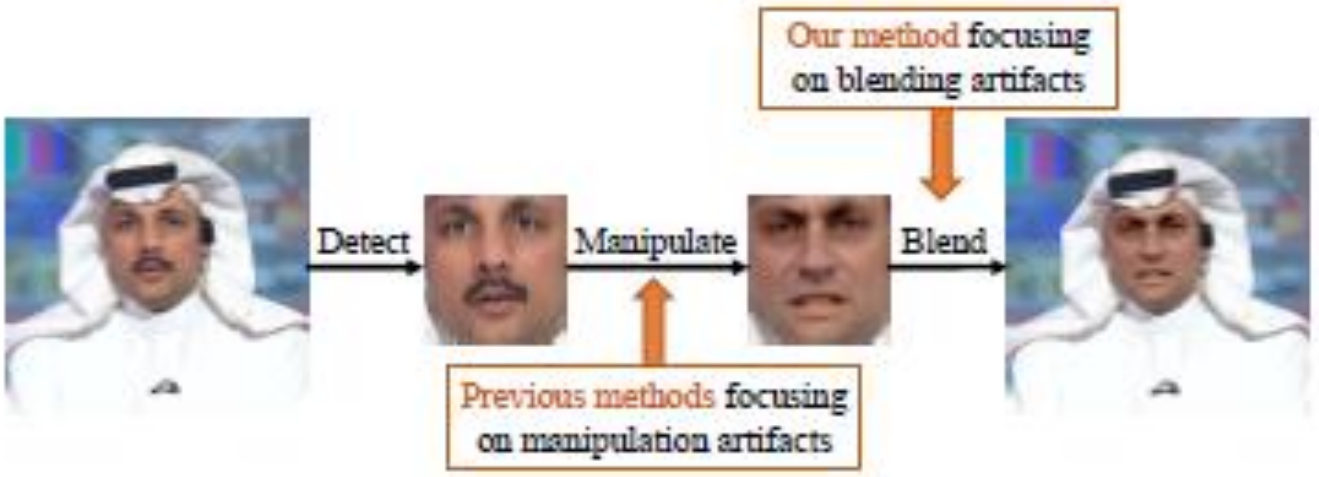
- Vidne napake na generiranih oz. spremenjenih slikah
  - Nepravilne zenice in odsev





Detekcija globokih ponaredkov (slike)

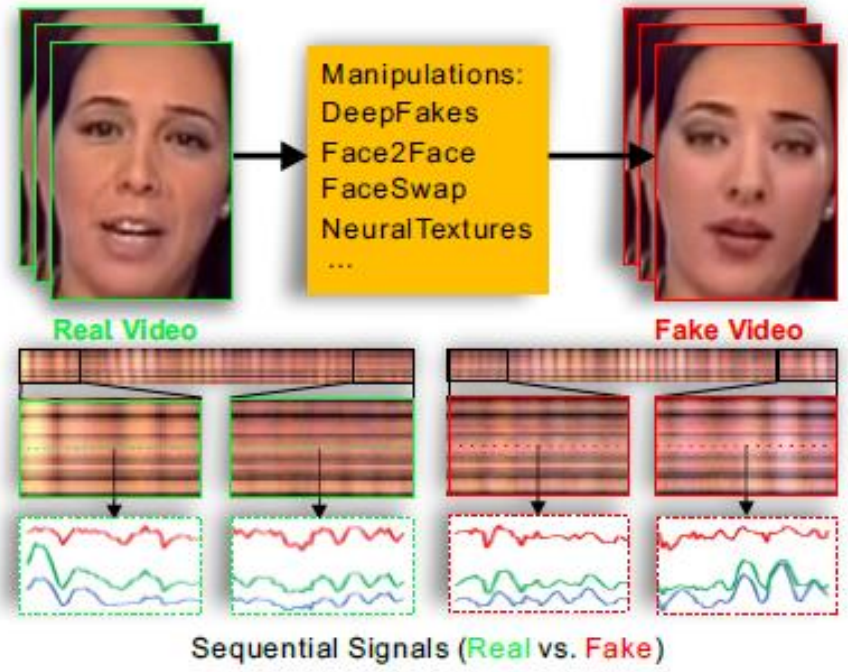
- Vidne napake na generiranih oz. spremenjenih slikah
  - Iskanje zameglitev na robovih



Li et al. 2020, Face X-ray for More General Face Forgery Detection

Detekcija globokih ponaredkov (**video**)

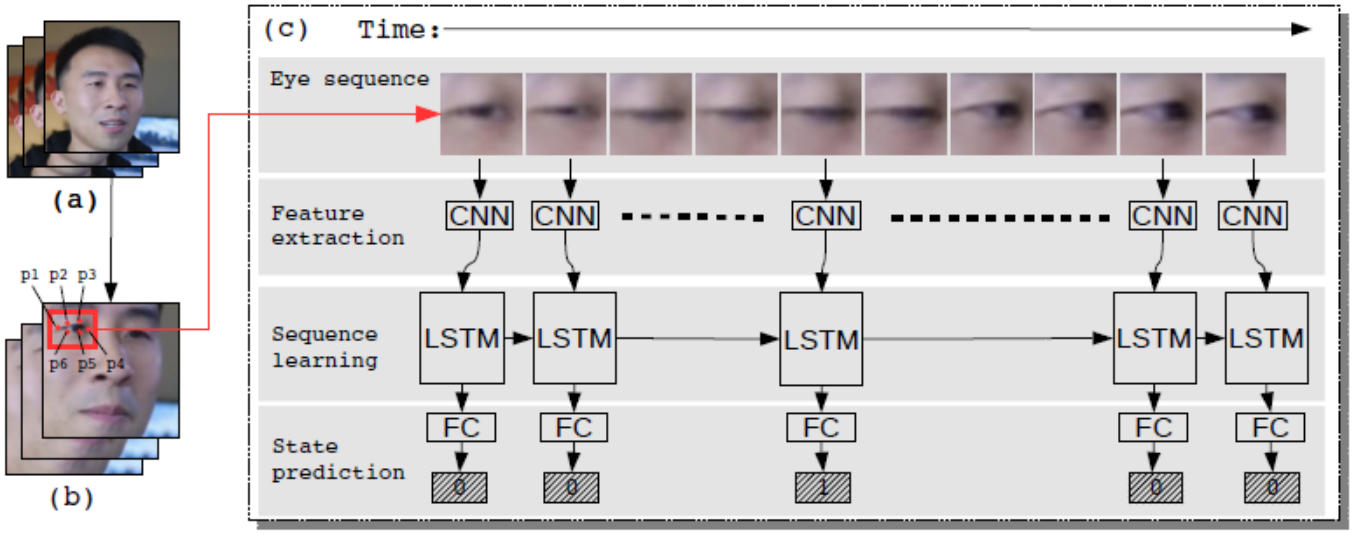
- Nepravilnosti med okvirji
  - Ritem utripanje srca (fotopletizmografija)



Qi et al. 2020, DeepRhythm: Exposing DeepFakes with Attentional Visual Heartbeat Rhythms

Detekcija globokih ponaredkov (video)

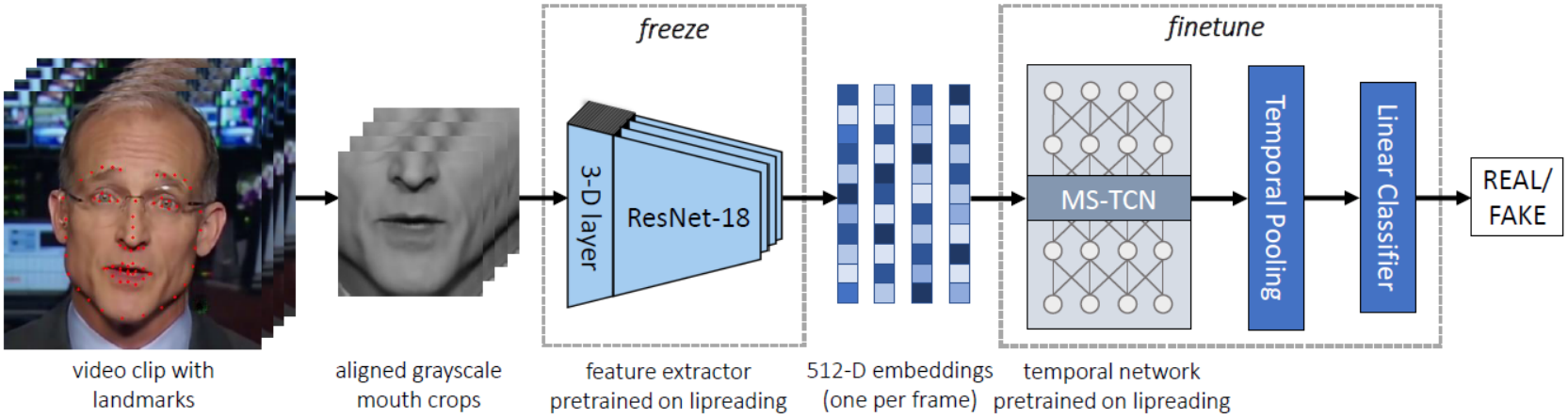
- Nepravilnosti med okvirji
  - Ritem utripanje srca (fotopletizmografija)
  - Neenakomerno mežikanje



Li et al. 2018, In Ictu Oculi: Exposing AI Generated Fake Face Videos by Detecting Eye Blinking

Detekcija globokih ponaredkov (video)

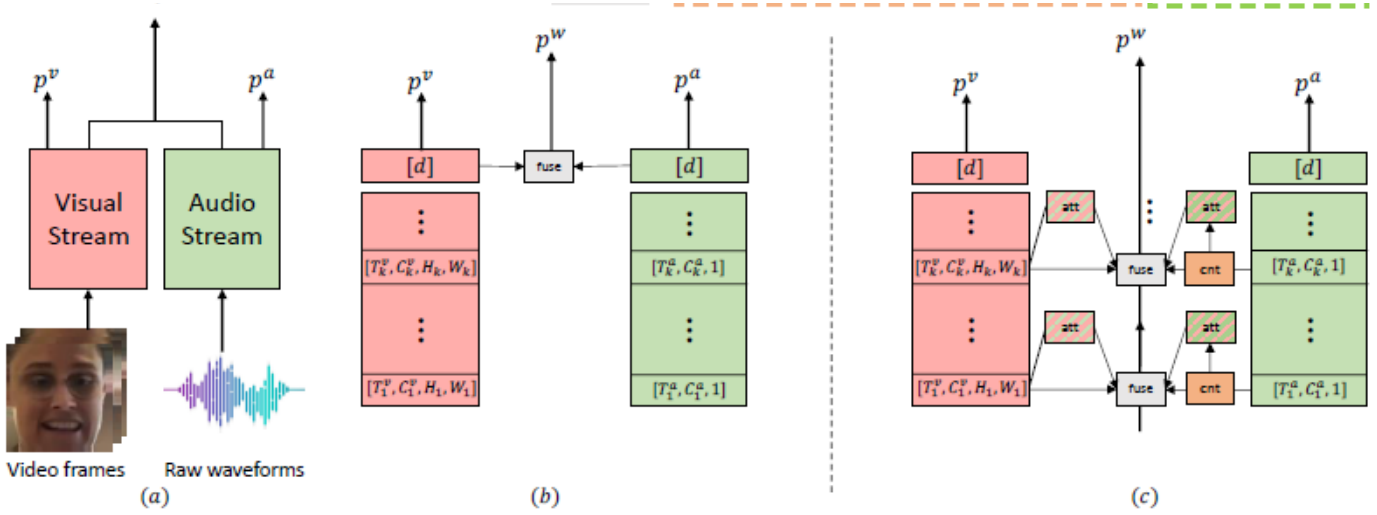
- Nepravilnosti med okvirji
  - Neskladnost v premikanje ustnic



Haliassos et al. 2021, Lips Don't Lie: A Generalisable and Robust Approach to Face Forgery Detection

Detekcija globokih ponaredkov (video)

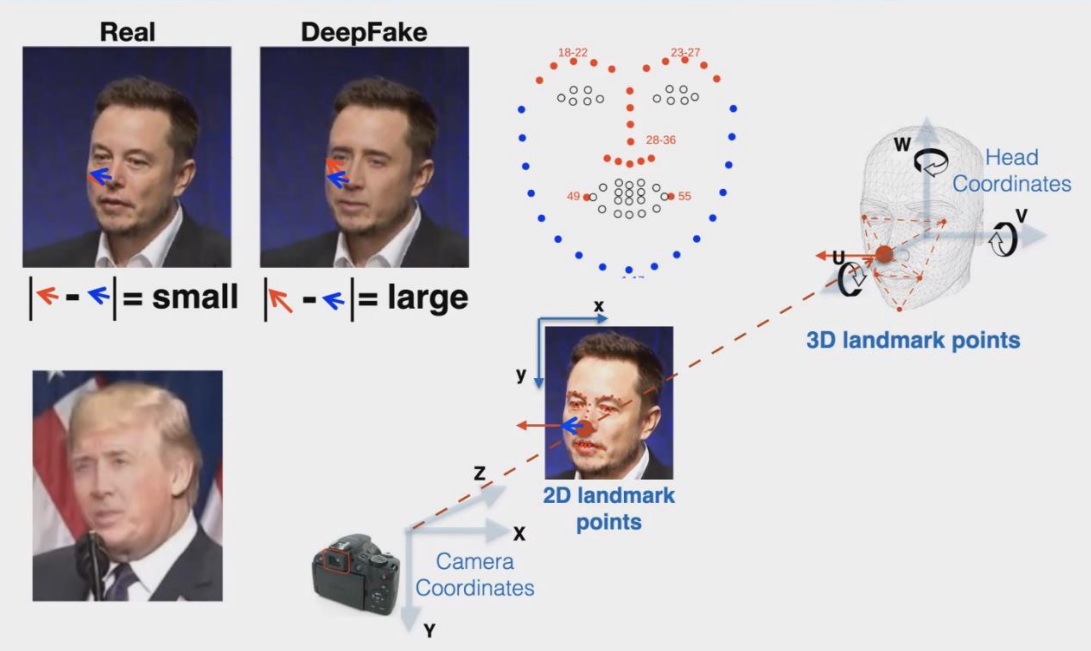
- Nepravilnosti med okvirji
  - Neskladnost med avdio in video



Zhou et al. 2021, Joint Audio-Visual Deepfake Detection

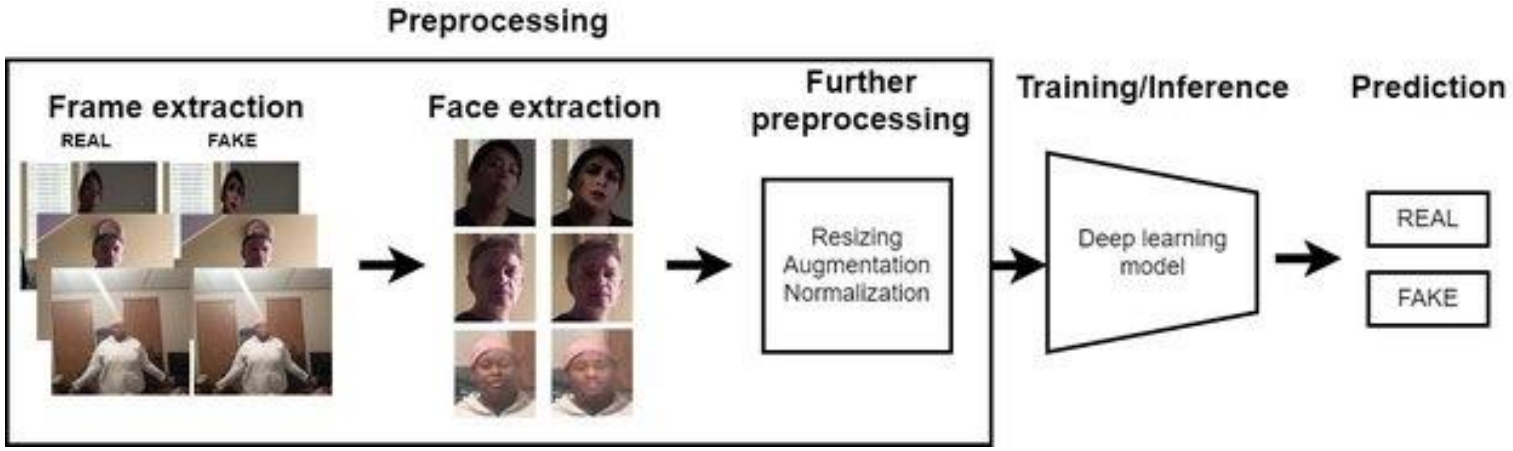
Detekcija globokih ponaredkov (video)

- Nepravilnosti med okvirji
  - Neskladni premiki
    - Nekonsistentnosti pri različnih položajih glave



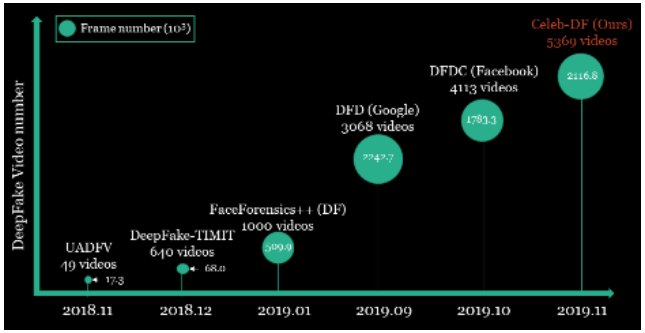
Yang et al. 2018, Exposing Deep Fakes Using Inconsistent Head Poses

CNN: Na podlagi podatkov



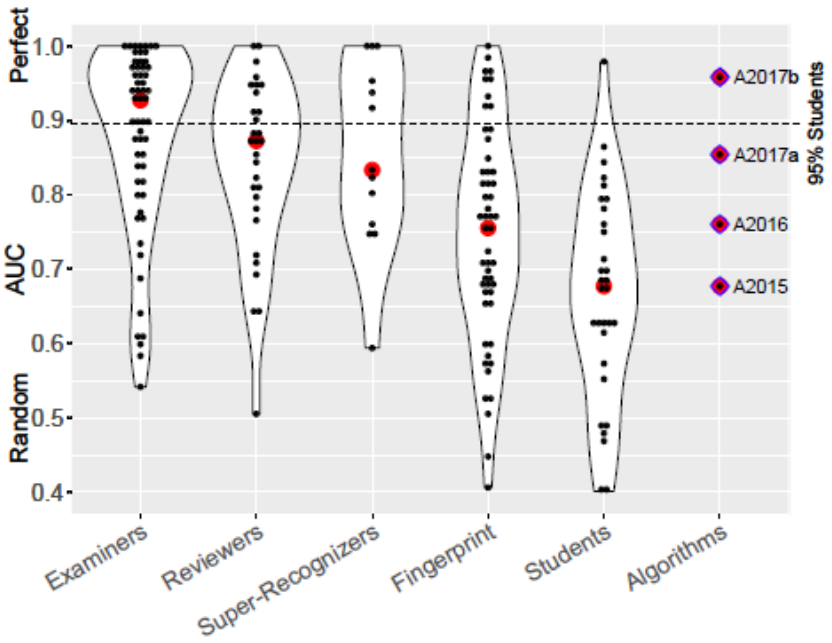
- Iskanje teksture
- Učenje na podlagi podslik (patch)
- Iskanje vzorca (Fingerprint) ki ga generira GAN

Baze


Frame number (10^3)
DeepFake Video number
Celeb-DF (Ours)
5369 videos
2116.8
DFDC (Facebook)
4113 videos
1781.1
DFD (Google)
3068 videos
724.7.7
FaceForensics++ (DF)
1000 videos
509.9
DeepFake-TIMIT
640 videos
68.0
UADFV
49 videos
17.3
2018.11
2018.12
2019.01
2019.09
2019.10
2019.11

Prepoznava obrazov

- Različne skupine
  - Forenzični preiskovalci za obrazno prepoznavo
  - Prepoznavalci obrazov
  - Superpoznavalci
  - Preiskovalci za prstne odtise
  - Študentje



Proti

- Računalniški sistemi (CNN)
  - A2015 – $A2017_2$

Vir: 1. P. Jonathon Phillips s sod., Face recognition accuracy of forensic examiners, superrecognizers, and face recognition algorithms
2. Ranjan, Rajeev, Carlos D. Castillo, and Rama Chellappa. "L2-constrained softmax loss for discriminative face verification."

University *of Ljubljana*
Faculty *of Computer and Information Science*

Hvala za pozornost

30. marec 2023

Lačen si full drugačen

Najdeš 11. napak?

Photoshop fails

Bi se radi preizkusili

Which face is real

- https://www.whichfaceisreal.com/

DETECT FAKES

- https://detectfakes.media.mit.edu/

Ali sta osebi isti?

Ali sta osebi isti?

Ali sta osebi isti?

Pornoindustrija

Politika

Dictators - Vladimir Putin.mp4

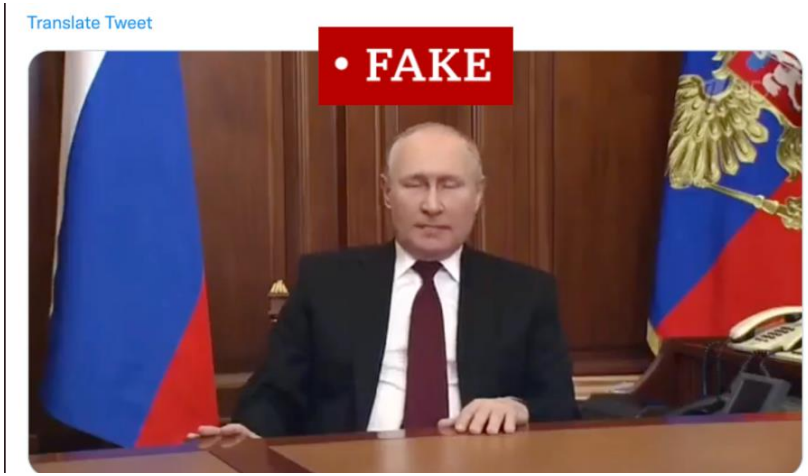
Dictators - Kim Jong-Un.mp4

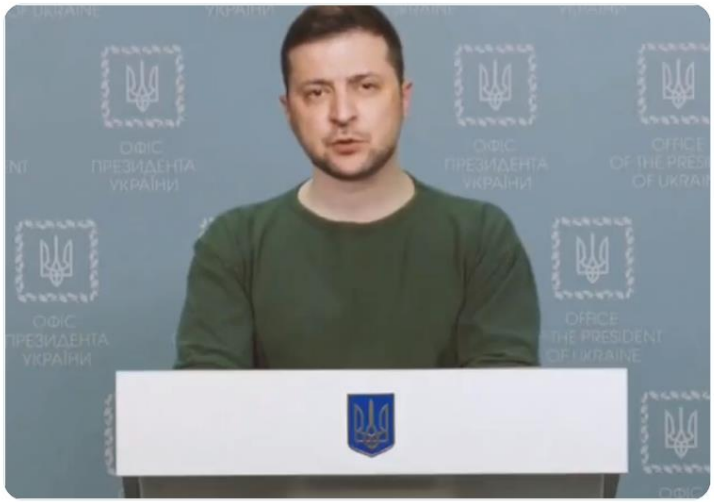
- Politična kampanija

https://www.karlsnotes.com/deepfakes-in-an-indian-election-campaign/

Politika

- Vojna v Ukrajini
  - Putin razglaša mir in Zelenski govori o predaji Rusiji





V splošnem pa tehnologija ni škodljiva

- Pomembno je kdo jo uporablja in za kakšen namen
- Vsi ponaredki pa niso škodljivi
  - Uporabimo za deidentifikacijo oseb, da zaščitimo identiteto
    - *Dobrodošli v Čečeniji (Welcome to Chechnya)*, ZDA, 2020
      They’re using artificial intelligence to overlay faces of volunteers on top of those of survivors.
  - V filmski industriji
    - Prikaz osebe v mlajših letih
    - Za izboljšanje obstoječih scen
    - Za nadomestitev igralcev
  - V industriji iger
  - V modni industriji

The future of gaming! Fifa 21 revamped using A.I. deepfa...

VFXChris Ume

The Irishman - Robert De Niro

Digital de-ageing of actors including **Robert De Niro** and Al Pacino.

YouTuber and prolific deepfaker Shamook places their own de-aged De Niro side-by-side with the original, and it's certainly impressive.

Vir:
https://www.youtube.com/watch?v=dHSTWepkp_M&t=1s

Dobrodošli v Čečeniji

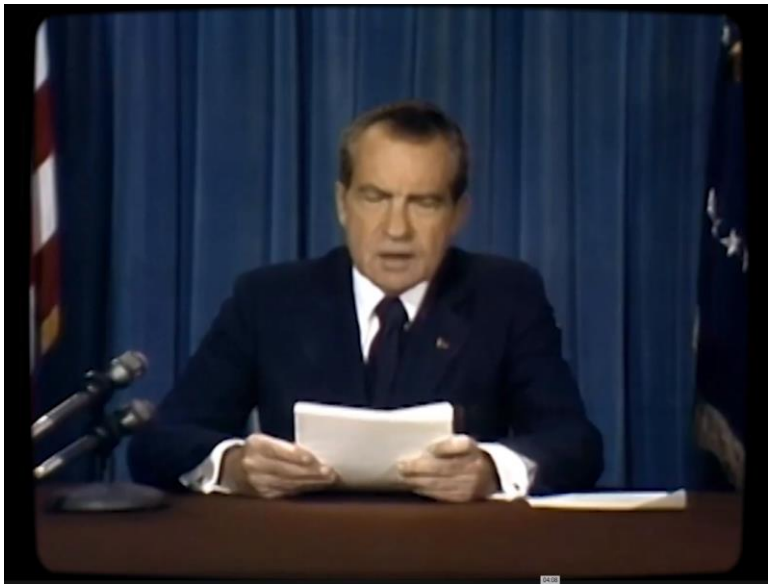

VFXCHRISUME.COM
Premier League
adidas
MANCHESTER UNITED
adidas
MANCHESTER UNITED

FIFA 2021?

Deepfake: The Irishman - Robert De Niro

"V primeru nesreče na Luni".

- Najbolj prepričjivi so posnetki, kjer ponaredimo zvok (glas) in video



20. julija 1969 sta Buzz Aldrin in Neil Armstrong pristala na Luni, a ta njuna pot, na kateri sta bila skupaj z Michaelom Collinsom, je bila polna nejasnosti, nevarnosti in obilice poguma.

Na Massachusettsovem inštitutu za tehnologijo (MIT) so pripravili ponaredek, ki kaže, kaj bi se zgodilo, če jima ne bi uspelo. Med drugim takratni predsednik **Richard Nixon** prebere "rezervni" govor.

Gre za resničen govor, ki ni bil nikoli prebran ali posnet, so ga pa leta pozneje našli v arhivih.

Vir: https://moondisaster.org/film