



Neumann János Egyetem
Műszaki és Informatikai
Kar

Képfavító algoritmusok hatékonyágának vizsgálata gépi látáson alapuló forgalomszámláláshoz

Szekeres György
Mérnökínformatikus hallgató, BSc
IA6OZQ

Témavezető:
Dr. Drenyovszki Rajmund,
adjunktus

2023

Tartalomjegyzék

TARTALOMJEGYZÉK	2
BEVEZETÉS	3
1. PROBLÉMAFELVETÉS.....	4
1.1. ÁLTALÁNOS PROBLÉMÁK	4
1.1.1. A képjavitás és a gépi látás általános kérdései	4
1.1.2. A járműfelismerés feladata	5
1.1.3. A képjavitó algoritmusok hatékonysága a járműfelismerésben	5
1.2. KÉPJAVÍTÓ ALGORITMUSOK CSOPORTOSÍTÁSA	5
2. KÉPJAVÍTÓ ALGORITMUSOK	8
2.1. FIZIKAI KÉPJAVÍTÓ ALGORITMUSOK	8
2.1.1. Zajcsökkentő algoritmusok	8
2.1.2. Kontrasztot manipuláló algoritmusok	12
2.1.3. Képfelbontást módosító algoritmusok	14
2.2. STATISZTIKAI KÉPJAVÍTÓ ALGORITMUSOK	17
2.2.1. Színkorrekciós algoritmusok	17
2.2.2. Textúrajavitó algoritmusok	17
2.3. MESTERSÉGES INTELLIGENCIA ALAPÚ KÉPJAVÍTÓ ALGORITMUSOK	18
2.3.1. ArbSR	18
2.3.2. DiffBIR	19
2.3.3. ESRGAN	20
2.3.4. HAT	21
2.3.5. HCFlow-SR	22
2.3.6. Real-ESRGAN	22
2.3.7. ruDALL-E	23
2.3.8. SRResCGAN	23
2.3.9. Swin2SR	24
2.3.10. SwinIR	25
3. KÉPJAVÍTÓ ALGORITMUSOK EREDMÉNYEI.....	27
3.1. HAGYOMÁNYOS ALGORITMUSOK	27
3.2. MESTERSÉGES INTELLIGENCIA ALAPÚ KÉPJAVÍTÓ ALGORITMUSOK EREDMÉNYE	29
3.3. A HAGYOMÁNYOS ÉS AI ALAPÚ ALGORITMUSOK KOMBINÁLÁSA	30
ÖSSZEFOGLALÁS	37
ÁBRAJEGYZÉK	38
IRODALOMJEGYZÉK	39

Bevezetés

A gépi látás alapú forgalomszámlálás egyre népszerűbb technika a forgalom megfigyelésére és elemzésére. A forgalomszámláló rendszerek jellemzően kamerákat használnak a forgalom megfigyelésére a jelentős csomópontokban vagy körforgalmakban. A kamerák által készített fényképekből objektumfelismerő algoritmusok segítségével származtatják az autók számát és sebességét.

A gépi látás alapú forgalomszámlálásnak számos előnye van a hagyományos forgalomszámlálási módszerekkel szemben. A hagyományos módszerek, pl. az induktív hurokdetektor vagy fotocellás forgalomszámlálás drága és telepítése időigényes. Ezzel szemben a gépi látás alapú forgalomszámlálás gyorsan és alacsony költséggel megvalósítható.

A gépi látás alapú forgalomszámlálásnak a hatékonysága azonban több tényezőtől függ, például a kamera minőségétől, a megvilágítástól és a forgalmi viszonyoktól. A rossz minőségű vagy kis méretű objektumokat tartalmazó képek értelmezése kihívást jelent az objektumfelismerő algoritmusok számára.

A kutatásom célja, hogy meghatározza a képjavító algoritmusok hatékonyságát a gépi látáson alapuló forgalomszámlálás területén. A képjavító algoritmusok javíthatják a képek minőségét azáltal, hogy csökkentik a zajt, javítják a kontrasztot vagy növelik a képfelbontást. A képmódosítás javíthatja az objektumdetektációs algoritmusok pontosságát és megbízhatóságát.

1. Problémafelvetés

1.1. Általános problémák

1.1.1. *A képjavítás és a gépi látás általános kérdései*

A képjavítás és a gépi látás során számos kihívással találkozhatunk. A képjavítással a cél a képek minőségének javítása, a zaj csökkentése, és a kontraszt növelése. A gépi látás célja a képek értelmezése és az objektumok azonosítása.

Ezek a területek közös kihívásokat jelentenek, többek között:

- Zaj: a zaj a képminőség romlásához vezethet, ami problémákat okozhat az objektumok felismerésében. A zajt különböző tényezők okozhatják, például a kamera érzékelőjének korlátjai, a rossz megvilágítás vagy a képek tömörítése.
- A kontraszt: a kontraszt hiánya azt eredményezheti, hogy a tárgyak kevésbé válnak láthatóvá. Az alacsony kontraszt gondot okozhat, mivel a különböző képterületek közötti fényerő különbség döntő szerepet játszik a tárgyak felismerésében.
- A felbontás: végül a felbontás is jelentős tényező, amelyet ebben a tekintetben figyelembe kell venni. A képfelbontás a képen belüli képpontok számát jelenti.

A zaj két típusba sorolható: színzaj és világossági zaj. [1] A világossági zajt két alkategóriára oszthatjuk megjelenési helye szerint. Lehet véletlenszerű, valamint helyhez kötött. A helyhez kötött világossági zajt másnéven sötétzajnak nevezzük. A sötétzaj fő okozója a szenzor magas hőmérséklete: minél melegebb az érzékelő, annál több lesz a sötétzaj a képen. A színzaj a képekben előforduló zaj másik típusa, amely kifejezetten a színes információk torzításához vezet. Ez a zaj színekkel kapcsolatos, és általában a különböző színárnyalatok, színtónusok és szaturációk helytelen megjelenítését eredményezi a képen. A színzaj gyakran kevésbé észrevehető, mint a világossági zaj, de bizonyos esetekben még mindig zavaró lehet, különösen, ha a képminőség romlik vagy a kontrasztos területeken jelenik meg. Míg a színes zaj hatékonyan kiküszöbölhető vagy szűrhető és javítható, addig a világoszaj szűrése már egy bonyolultabb folyamat, mivel minden típusú zajnak az eltávolításával vagy csökkentésével romlik a kép minősége oly módon, hogy az éles részletek nagyon elmosódnak, és egy életlen, homályos képet kaphatunk eredményül. [2][3]

1.1.2. A járműfelismerés feladata

A járműfelismerés a gépi látás egy speciális területe, amely a járművek képeken történő felismerésével foglalkozik. Többféle kihívást is magába foglal, többek között:

- Járművek mérete: az autók nagysága ingadozhat, ami miatt nehéz megkülönböztetni őket. A gépjárművek méretei eltérhetnek az autó fajtája, a forgalmi szituációk és a kamera távolsága alapján.
- Járművek formája: a járművek formája eltérő lehet, ami potenciálisan kihívást jelenthet az azonosításban. A járművek jellege függhet az autótípustól, a modelltől, a kamera optikájának típusától és a forgalmi körülményektől.
- Járművek környezete: a járműveket a környező tárgyak jelenléte miatt nehéz lehet azonosítani a képeken. A forgalmi viszonyok, az időjárás és a kamera távolsága mind befolyásolják a jármű környezetének megjelenését.

1.1.3. A képjavító algoritmusok hatékonysága a járműfelismerésben

A képjavító algoritmusok növelhetik a járműfelismerő algoritmusok hatékonyságát. A zaj csökkentése, a kontraszt fokozása és a képfelbontás növelése révén a képjavítás képes javítani az objektumfelismerés pontosságát és megbízhatóságát.

Ez különösen fontos azokban az esetekben, amikor a képen lévő zaj miatt a járműfelismerő algoritmusok hibázhatnak, és tévesen másnak azonosíthatnak egy tárgyat. A képek alacsony kontrasztja miatt a járművek nehezen különböztethetőek meg a háttértől, az alacsony képfelbontás pedig azt eredményezheti, hogy a járműfelismerő algoritmus nem veszi észre az apró részleteket.

A képjavítási technikák segíthetnek megoldani az előbbi problémákat, és javíthatják a járműfelismerő algoritmusok hatékonyságát.

1.2. Képjavító algoritmusok csoportosítása

A képjavító algoritmusok többféleképpen osztályozhatók, többek között:

- Működés szerint:
 - Fizikai képjavító algoritmusok,
 - Statisztikai képjavító algoritmusok,
 - Deep learning alapú képjavító algoritmusok.
- Alkalmazási terület szerint:
 - Képfeldolgozás,
 - Képrekonstrukció,

- Képtömörítés.
- Számítási igény szerint:
 - Alacsony,
 - Magas.
- Adatszükséglet szerint:
 - Adatfüggetlen,
 - Adatfüggő.

Fizikai képjavító algoritmusok közvetlenül a kép fizikai paramétereit, például zajt, kontrasztot vagy felbontást módosítják, amíg a statisztikai képjavító algoritmusok a kép általános tulajdonságait, mint például színkontraszt vagy textúrák, statisztikai szempontok alapján javítják. A fizikai algoritmusok a pixelértékeket befolyásolják, míg a statisztikai algoritmusok globálisabb módosításokat hajtanak végre. Mindkét típusnak megvannak a saját alkalmazási területei a képfeldolgozásban, attól függően, hogy melyik jellemzőt kívánják optimalizálni. A gépi tanulás területéről a mélytanuláson alapuló algoritmusok neurális hálózatok segítségével olyan kombinált algoritmust alkotnak, amely egyszerre képes módosítani a fizikai és a statisztikai tulajdonságokat.

Olyan helyzetekben, amikor a kép sérült vagy hiányos, képrekonstrukciós algoritmusokat alkalmaznak. Ha a sérülés mértéke minimális, a hiányzó képpontok kiszámítása például a szomszédos képpontok matematikai átlagából lehetséges. Olyan esetekben azonban, amikor a sérülés jelentős, mesterséges intelligencia technikák alkalmazására lehet szükség.

A képtömörítési algoritmusok különböző technikákat alkalmaznak a fájl méret csökkentésére és ennek következtében a tárhely megtakarítására, amivel a kép minőségét is megváltoztatják, gyakran rontják. Az algoritmusok erőforrásigényét több tényező befolyásolja, például az algoritmus bonyolultsága, a számítási mennyiség és a kép mérete. Gyakori tévhit, hogy egy kép felbontását full HD vagy 4K felbontásúnak minősítik. Ez azonban helytelen, mivel a full HD és a 4K a képben lévő képpontok száma alapján kerül meghatározásra. A képfelbontást dpi-ben, azaz pont per hüvelykben mérik, amit pixelsűrűségnek is neveznek. A képpontsűrűség mérésének másik módja a pont/mm-ben történő mérés. Egyszerűbben fogalmazva, ez a kép mm-enkénti pontok számát jelenti. Például egy 300 dpi-s vízszintes A/4-es lap 3508x2480 képpontnak felel meg, míg ugyanez a lap 600 dpi-s felbontásnál magasabb, 7016x4961 pixelszámot eredményez. Mint látható, a kép fizikai mérete nem változik, de a képet alkotó pixelek száma nő. [4]

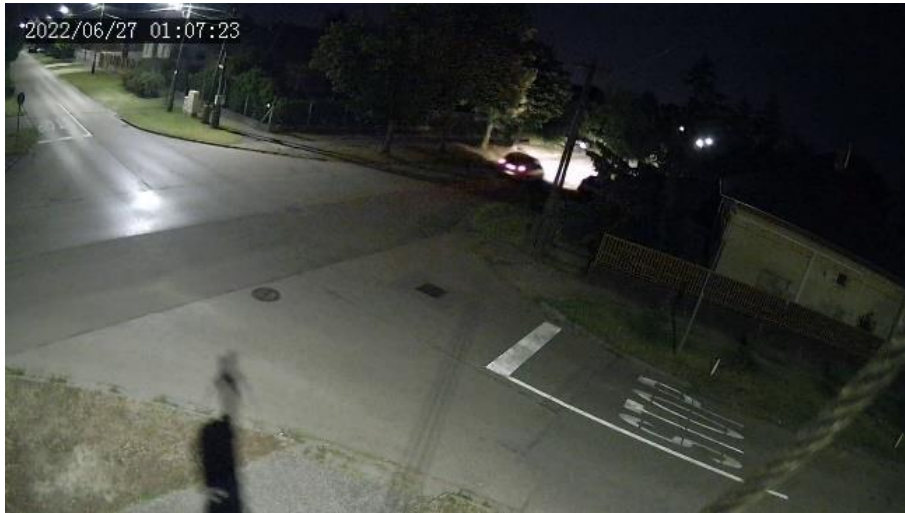
Fontos megjegyezni, hogy egy algoritmus lehet adatfüggetlen vagy adatfüggő. Az adatfüggetlen algoritmusok kis mennyiségű adatra vagy képre támaszkodnak az algoritmus betanításához, míg az adatfüggő algoritmusoknak nagy adathalmazokra van szükségük a folyamat elvégzéséhez.

Dolgozatom célja, hogy áttekintést nyújtson a különböző képjavító algoritmusokról, valamint azok funkcionális osztályozáson alapuló képminőségre gyakorolt hatásáról.

2. Képjavító algoritmusok

2.1. Fizikai képjavító algoritmusok

A kép fizikai tulajdonságait módosító algoritmusokat soroljuk ebbe a kategóriába. Ezek alapján ide tartoznak a zajt, kontrasztot és felbontást manipuláló algoritmusok. Ezeket az algoritmusokat az 1. ábrán látható képen fogom bemutatni.



1. ábra: Alapkép

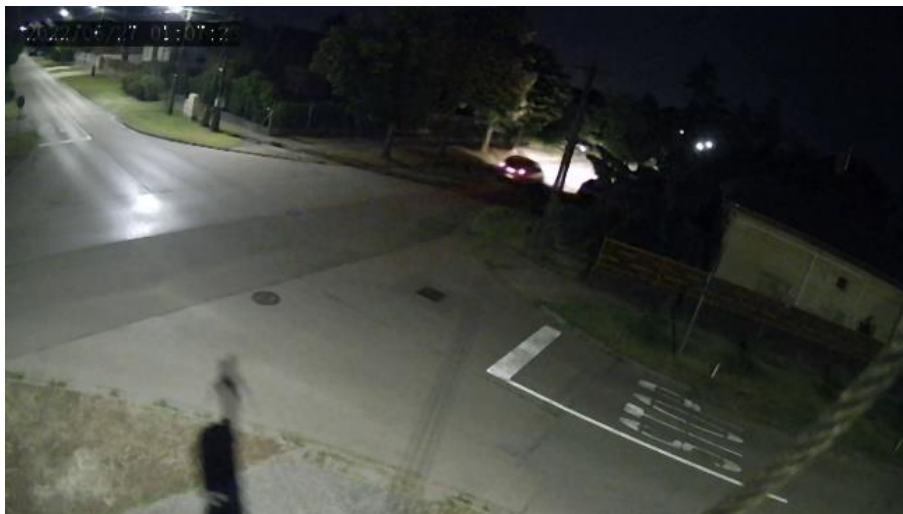
(Forrás: saját kép)

2.1.1. Zajcsökkentő algoritmusok

A zajcsökkentő algoritmusok a gépi látás egyik fontos területét képezik. A zaj a kép minőségének romlását okozhatja, például a képet homályossá vagy zavarossá teheti. A zajcsökkentő algoritmusok segítenek csökkenteni a zajt, így javítva a kép minőségét és a gépi látási algoritmusok teljesítményét. A zajcsökkentő algoritmusok általában a kép fizikai tulajdonságait módosítják, például a zaj frekvenciáját vagy a zajmintázatát. Ezek az algoritmusok általában egyszerűek és gyorsak, de nem mindig hatékonyak. [5][6]

A leggyakoribb zajcsökkentő algoritmusok a következők:

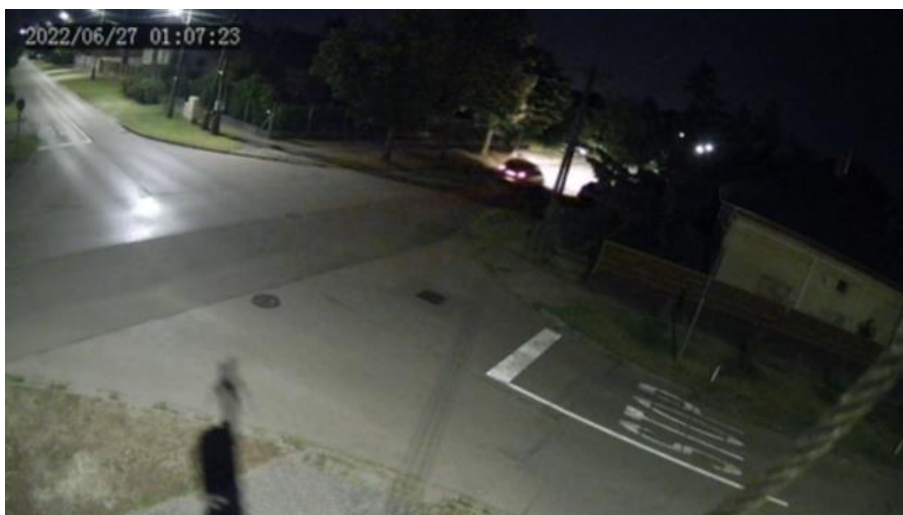
- Medián-szűrő: a kép minden pixeléhez a legközelebbi néhány pixel mediánját veszi. Ez az algoritmus hatékonyan csökkenti a zajt, de elmoshatja a kép élességét.



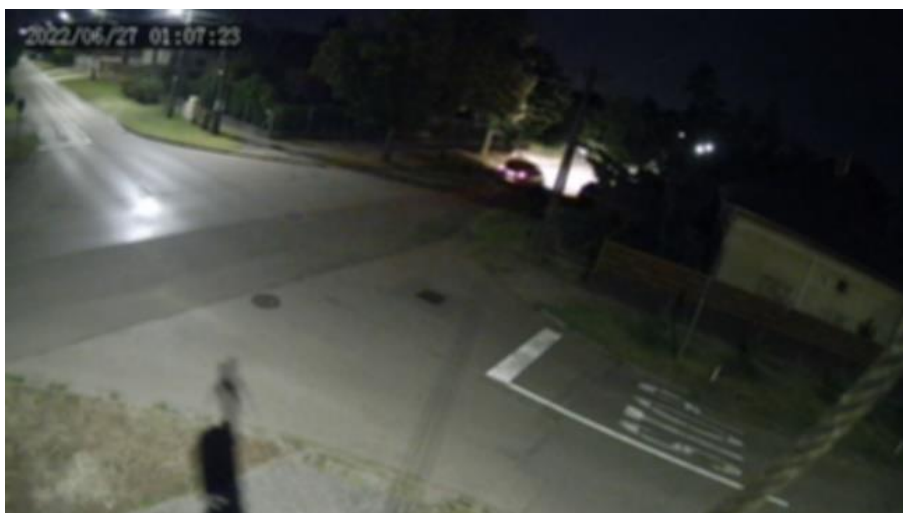
2. ábra: Medián-szűrő alkalmazása

A 2. ábrán látható, hogy ugyanaz algoritmus közepes hatékonysággal csökkentette a zajt a forrásképen (1. ábra), de egyben az eddig jól kivehető apróbb részleteket elmosta, ezáltal megnehezítve a különböző objektumok detektálását a képen. [6][7]

- Gauss-szűrő: Ez az algoritmus a kép minden pixeléhez egy Gauss-függvényt alkalmaz. Hatékonyan csökkenti a zajt, és kevésbé mosódik el a kép, mint a medián-szűrő esetében.



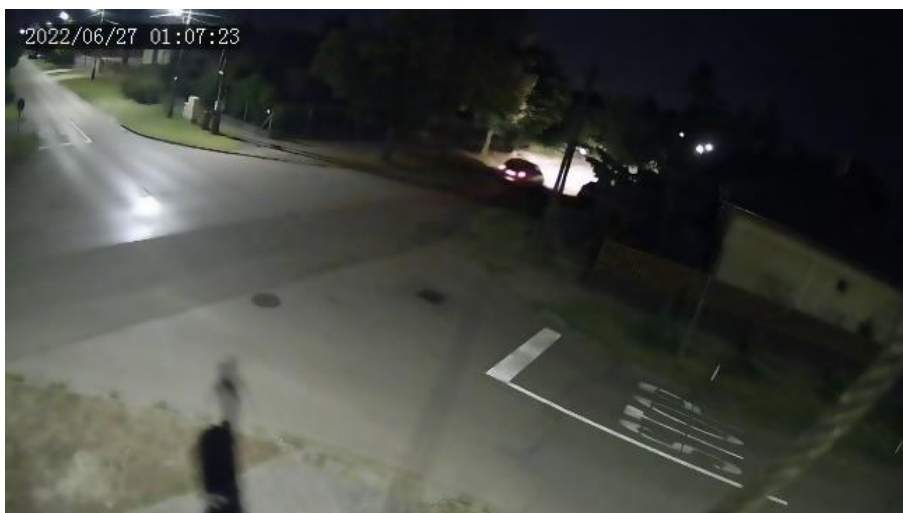
3. ábra: Gauss-szűrő alkalmazása 5-ös méretű kernellel



4. ábra: Gauss-szűrő alkalmazása 9-es méretű kernellel

A Gauss-szűrő hatása a 3. és 4. ábrán figyelhető meg, és ennél az algoritmusnál kernel méretet kell megadni, amit 3, 5, 9 értékekre választottam. Ez a mérőszám megadja a súlymátrix méretét, hogy mekkora területet vegyen figyelembe, valamint milyen mértékben figyelje a zajt. A kernelszám növelésével javul a zajszűrés minősége, de ezzel arányosan csökken az élessége is, ezáltal egy elmosódott képet kapunk végeredményül. [6][8]

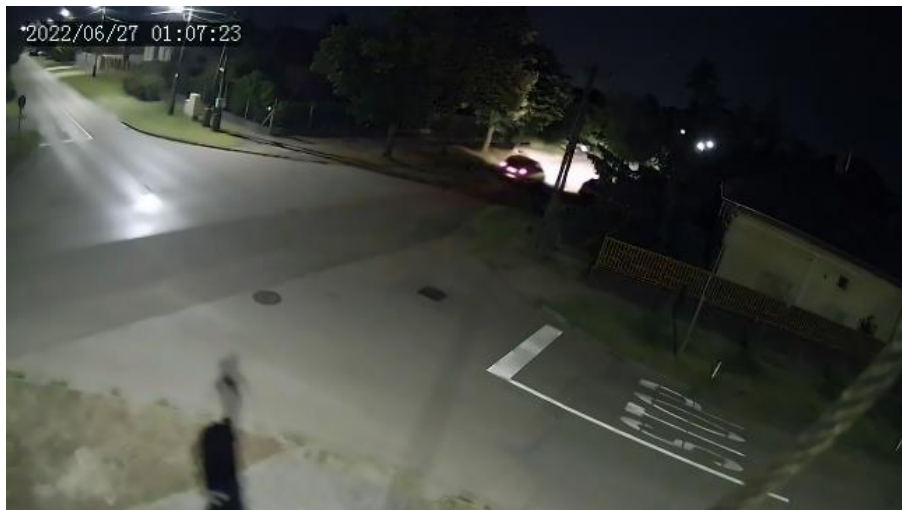
- Bilaterális szűrő: Ez az algoritmus a kép minden pixeléhez egy közeli szomszédokból származó súlyozott átlagot veszi. Ez az algoritmus hatékonyan csökkenti a zajt, miközben megtartja a kép élességét.



5. ábra: Bilaterális szűrő alkalmazása

A bilaterális szűrő alkalmazása (5. ábra) után az apróbb részletek is nagyobb eséllyel detektálhatóak, és a zajszint is csökkentésre került. A nagy kontraszttal rendelkező objektumok jobban kitűnnek a képből, ezáltal is segítve a detektálást. [9][10]

- Non-Local Means (NLM) szűrő: a kép mintázatát vizsgálja, és minden pixelhez egy hasonló mintázatból származó képpontok súlyozott átlagát rendeli hozzá. Ez az algoritmus hatékonyan szünteti meg a zajt, miközben megtartja a kép élességét, és részletgazdagságát.



6. ábra: Non-local means szűrő alkalmazása

Az NLM szűrő (6. ábra) alkalmazása után a zaj szinte teljesen eltüntetésre került, de a kép a minőségét megtartotta. A sötét részeken lévő fák apróbb részletei is jól kivehetőek, tehát az eddig bemutatott algoritmusok közül ez vezetett a legjobb eredményhez a kép részleteinek rontása nélkül. [9][10]

Mint ahogyan az ábrákon is megfigyelhető (2-6. ábra), a szűrő algoritmusok képesek hatékonyan csökkenteni a zajt, ezáltal javítva a kép minőségét. Az algoritmusok gyorsan, nagy erőforrásigény nélkül futtathatóak. Jól megfigyelhető a képeken, hogy nem mindig voltak hatékonyak, mivel szinte minden esetben elmosták a képek élességét, ezáltal homályossá téve azt.

A zajcsökkentő algoritmusoknak fontos szerepe van a gépi látásban, mivel bár korlátjaik közé tartozik az élesség elmosása, valamint a hatékonyság, de ezeket figyelembe véve kell kiválasztani azt az algoritmust, amely megfelel az adott feladatnak. Fontos azonban azt megjegyezni, hogy jelen esetben éjszakai képeken mutattam be ezeknek az eredményét, de teljesen eltérő eredményt is kaphatunk, ha más szögben áll a kamera, vagy más szögben érkeznek a fénysugarak. Ezeket az algoritmusokat a gyakorlatban előfeldolgozásra szokás használni.

Az algoritmusok fejlesztésén jelenleg is sokan dolgoznak, hiszen fontos szerepe van a gépi látásban. A modernebb zajszűrők már mesterséges intelligenciával vannak felvértezve, ezáltal sokkal jobb eredményeket elérve az élek elmosódása nélkül is.

2.1.2. Kontrasztot manipuláló algoritmusok

A kontrasztot módosító fizikai képjavító algoritmusok a kép kontrasztját változtatják meg. A kontraszt a kép egyes területeinek fényerősségének különbsége. Alacsony kontraszt esetén az objektumok kevésbé jól láthatók. Ezeknek az algoritmusoknak két fő csoportja létezik:

- fényerősséget szabályozó algoritmusok és
- gammakorrektációs algoritmusok.

A fényerősséget szabályozó algoritmusok a kép fényerősségét változtatják meg a paraméterezés függvényében. Ennek az egyik legegyszerűbb módja, ha a pixelek értékét egy konstans számmal szorozzuk meg vagy osztjuk el. Ennek vizsgálatára két algoritmust próbáltam ki. Az algoritmusokat a korábban bemutatott képen futtattam le (1. ábra), és mind a kettő a fényerősséget növeli. Az eredmény jól megfigyelhető a következő ábrákon:



7. ábra: Fényerősség növelése I.



8. ábra: Fényerősség növelése II.

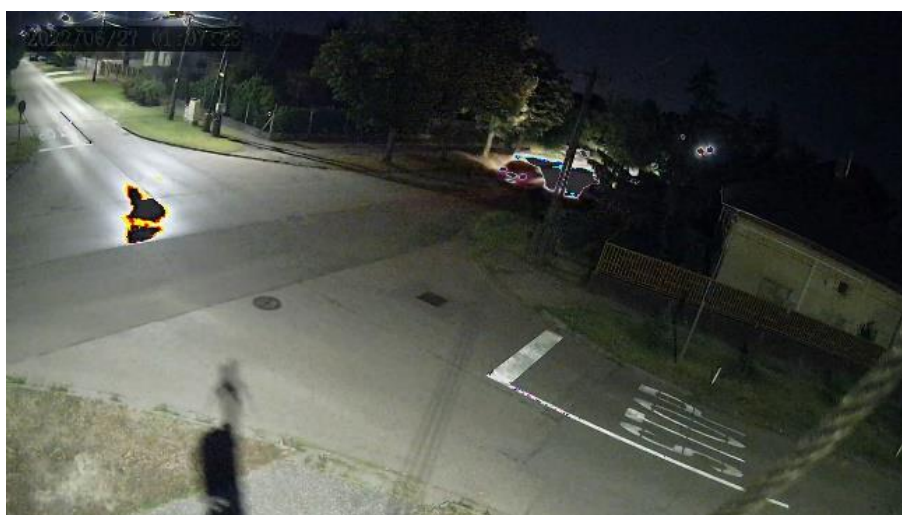
A fenti ábrákon (7-8. ábra) jól látható, hogy a fényerő növelésével jobban kivehetővé váltak a tárgyak a képen.

A gammakorrekciós algoritmusok a kép fényerősségének eloszlását változtatják meg, ezáltal javítva a részletek jobb láthatóságát. A gammakorrekció egy lineáris transzformáció, amely a pixelek értékét egy másik tartományba helyezi át. Két ilyen algoritmust teszteltem, az egyik egy gammakorrekciós algoritmus, a másik pedig egy kontrasztosztás algoritmus. [11][12]



9. ábra: Gammakorrekciós algoritmus

A fentebbi ábrán (9. ábra) jól látható, hogy az eddigi algoritmusok közül ez növelte a legeredményesebben a fényerőt. Az autó is nagyon jól kivehető a képről, valamint a távolabbi tereptárgyak is láthatóvá váltak, bár ez a jelenlegi problémánál nem fontos.



10. ábra: Kontrasztosztás algoritmus

A kontrasztosztás algoritmus (10. ábra) megfigyelhető, hogy az algoritmus a legvilágosabb és a legsötétebb helyeken egy színes határral körülvett fekete foltot hoz létre, ezáltal megsemmisítve az esetlegesen fontos részleteket. Jelen képen az algoritmust 1,2-es

értékkel futtattam, az ajánlott 2,0 helyett, mivel az utóbbi érték esetében nagy felületet takart volna le az algoritmus.

A kontrasztot befolyásoló algoritmusok elsősorban a jobb detektálást hivatottak segíteni, de ahogyan a zajszűrő algoritmusok között, itt is található olyan algoritmus, amely rontja a felismerés esélyeit. Ezekről az algoritmusokról elmondhatóak, hogy egyszerű működésűek, gyorsan lefutnak és nem igényelnek nagy erőforrásokat. Azonban arról sem szabad megfeledkezni, hogy eltorzíthatják a kép eredeti megjelenését, valamint a hatékonyságuk sem mindig megfelelő.

2.1.3. Képfelbontást módosító algoritmusok

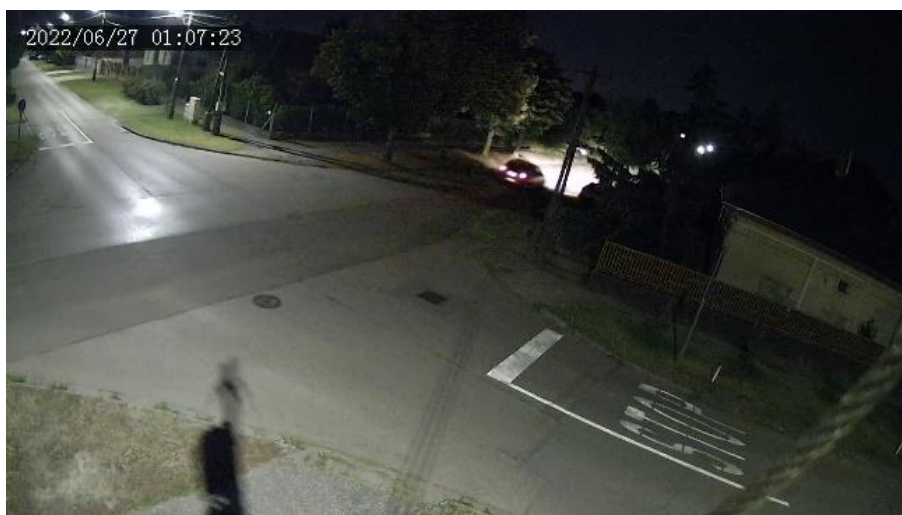
A képfelbontás módosítását három alapvető kategóriába sorolhatjuk. Az első kategória az "upscale" vagyis felskálázás, amely a kép felbontását növeli, ezzel javítva az objektumdetektációs algoritmusok teljesítményét.

A második alkalmazási lehetőség a "downscale" vagyis a leskálázás, ahol a kép felbontását csökkentjük. Ezt különösen olyan esetekben használjuk előnyösen, ahol a kép felbontása magas, és a képminőség csökkenése elfogadható. A downscale fő előnye, hogy jelentősen csökkenti az erőforrásigényt és felgyorsítja a feldolgozást, mivel a kevesebb adat könnyebben kezelhető.

A harmadik lehetőség a képtömörítés, amely a kép felbontásának módosításával kisebb méretű fájlokat eredményez. A képtömörítés hatékony módszer a tárolási és átviteli célokra, mivel csökkenti a fájl méretét, miközben a vizuális tartalom megmarad.

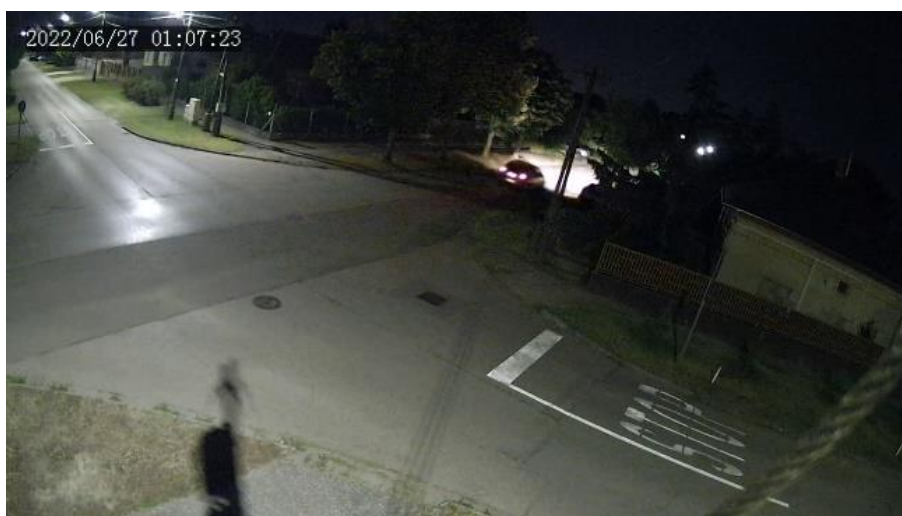
Négyféle algoritmuson keresztül vizsgáltam a nem mesterséges intelligencia alapú felbontásmódosító algoritmusokat, melyek a következők:

- interpoláció,
- decimation,
- bilineáris interpoláció és
- közelítő interpoláció.



11. ábra: Interpoláció kétszeres nagyításban

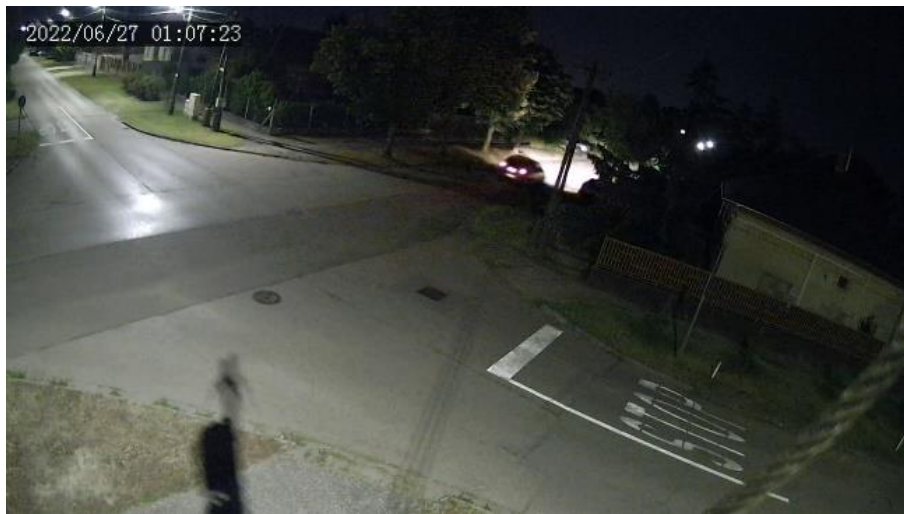
A kétszeres nagyítás interpolációs eljárással, bár a kép felbontása megnő, általában nem jár javulással a kép minőségében. Az interpolációs algoritmusok új képpontokat hoznak létre a meglévők között, melyek színét a szomszédos képpontok alapján számolják ki. Ennek eredményeként a kép gyakran "kinyújtott" hatást kelt, ami a minőség romlásához vezethet. Az interpolációval létrehozott képek általában homályosabbak és kevésbé részletgazdagok lehetnek a szomszédos képpontok összemosódása miatt.



12. ábra: Decimation algoritmus félszeres kicsinyítés

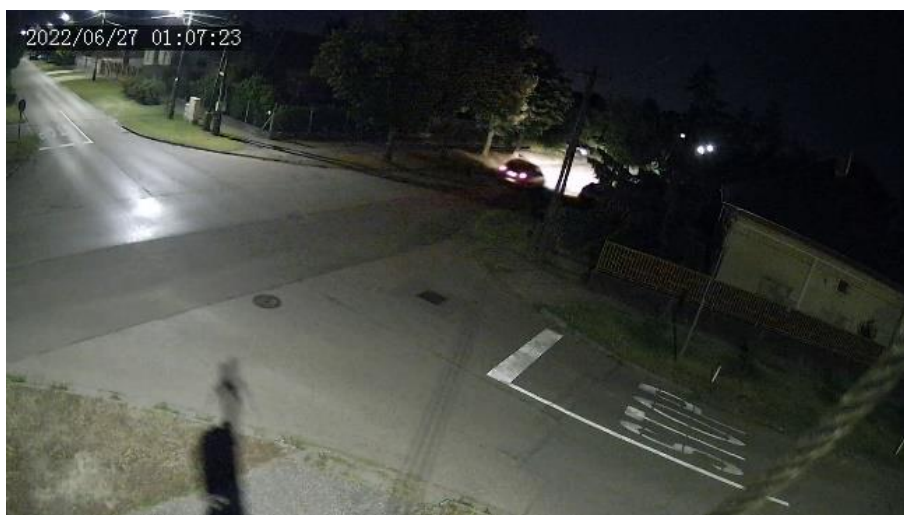
A decimation algoritmus (12. ábra), bár eredetileg a képek kicsinyítésére szolgál, lehetőséget nyújt a kép nagyítására is, ha a skálázási tényező 0 és 1 közötti érték. A képminőséget tekintve az algoritmus javítást hozhat, amelyet annak köszönhetünk, hogy a decimation először csökkenti a kép méretét, majd megnöveli azt a kívánt méretre, közben zajszűrést alkalmazva. Ennek következtében a képminőség javul, és a kép éle is simábbá válik. Ezen előnyök mellett a decimation további előnye, hogy a nagyított képekben a sötétebb területeken kevésbé jelentkező zaj jellemző. Például 5000%-os nagyítás esetén is

megfigyelhető a simább él, és a sötét területek zajszintje jelentősen alacsonyabb, mint az eredeti képen. [13][14]



13. ábra: Bilineáris interpoláció kétszeres nagyításban

A bilineáris interpolációs algoritmus (13. ábra) egy elterjedt módszer a képfeldolgozásban, amely az új képpontok létrehozásához négy szomszédos képpontot vesz figyelembe. Ezen szomszédos pontok értékei alapján a bilineáris interpoláció simább és lágyabb átmeneteket hoz létre az új képpontok között. Ennek eredményeként az interpolált kép minősége általában javul az eredetihez képest, és a kemény élsimítások enyhülnek. Az előnyök és az árnyoldalak ismertek: a kép minősége javul, de az élsimítás hatása is nő, amely lehetővé teszi a zaj kismértékű csökkentését. [15]



14. ábra: Közelítő interpoláció kétszeresen nagyítva

A közelítő interpoláció (14. ábra) egy olyan elterjedt képfeldolgozási eljárás, amely az új képpontok létrehozásához a legközelebbi szomszédos képpontokat használja. Ez a technika gyakran alkalmazható a képek méretének csökkentésére, amit kicsinyítésnek nevezünk. Nagyításkor azonban hátrányai is vannak: az eredmény kép pixeles és zajos lehet.

Ennek oka, hogy a közelítő interpoláció egyszerűen másolja az eredeti kép pontjait, ami az eredeti kép részletességének elvesztését eredményezheti, és a zajok fokozódását okozhatja a képen. [13][14]

A képfelbontás-módosító algoritmusok olyan kulcsfontosságú eszközök a gépi látásban, amelyek lehetővé teszik a rendszerek számára, hogy hatékonyan dolgozzanak nagy mennyiségű vizuális adattal. Ezek az algoritmusok a képek felbontását, méretét és minőségét módosítják, ami segít a gépi látási alkalmazásoknak javítani az adatfeldolgozás hatékonyságát. A módosított képek lehetővé teszik a gépi látás rendszerek számára, hogy gyorsabban és hatékonyabban végezzenek feladatokat, például objektumfelismerést vagy kép osztályozást. Ennek révén a gépi látási alkalmazások jobban teljesítenek, és kevésbé terhelik a számítási erőforrásokat, amelyek kulcsfontosságúak az alkalmazások hatékony működése szempontjából.

2.2. Statisztikai képjavító algoritmusok

A statisztikai képjavító algoritmusok olyan komplex technikák, amelyek a képek minőségének növelésére összpontosítanak a képek statisztikai tulajdonságainak finomhangolásával. Ezen algoritmusok általában magasabb számítási igényűek és lassabbak lehetnek a fizikai képjavító módszereknél, azonban rendszerint hatékonyabbak is.

Ezeket a statisztikai képjavító algoritmusokat két fő kategóriába lehet sorolni:

- színekorrekciós algoritmusok,
- textúrajavító algoritmusok.

2.2.1. Színekorrekciós algoritmusok

Ezek a megoldások a képek színegyensúlyát, kontrasztját és színhőmérsékletét módosítják azzal a céllal, hogy a képek színei élénkebbek legyenek, a színhelyes kiegyensúlyozás pedig javított legyen. Gyakran alkalmazzák a digitális fényképezés és a grafikai tervezés területén annak érdekében, hogy a képek jobban reprezentálják az eredeti látványt.

2.2.2. Textúrajavító algoritmusok

Ezek az eljárások a képek textúráinak minőségét és részletességét növelik. Céljuk a képek élességének javítása, a zaj csökkentése és a textúrák részletes bemutatása. Ezek az algoritmusok gyakran használhatók az orvosi képfeldolgozásban, a képfelismerésben és a távérzékelésben a képi információk hatékonyabbá tételére.

Mindkét csoportban alkalmazott algoritmusok közös vonása, hogy többek között hisztogram-kiegyenlítést és gamma korrekciót alkalmaznak. Annak ellenére, hogy a hisztogram-kiegyenlítés hatékony lehet, szembetűnik, hogy kizárólag 1:1 képarányú képeken működik megfelelően, és ezért fölösleges lépéseket hozhat létre a feldolgozási folyamatban, valamint az erőforrásigényt is növeli, ami jelentős költségeket okozhat.

2.3. Mesterséges intelligencia alapú képjavító algoritmusok

Mesterséges intelligencia alapú képjavító algoritmusok, vagy AI-alapú képjavító eljárások, az intelligens gépi tanulás és adatfeldolgozás terén nyújtanak megoldásokat a képek minőségének és tartalmának javítására. Ezek az algoritmusok olyan kifinomult technikákat alkalmaznak, amelyek lehetővé teszik a zaj kiszűrését, a részletek kiemelését, a színekorrekciót és az élesség növelését, hogy a végeredmény megegyezzen vagy akár javítsa az eredeti képet.

Az AI-alapú képjavító algoritmusok lényegesen különböznek a hagyományos módszerektől, mivel azok tanulnak az adatokból és adaptálódnak az adott képekhez. A gépi tanulási és neurális hálózatok használata lehetővé teszi számukra, hogy komplexebb mintázatokat és összefüggéseket ismerjenek fel a képeken, ami eredményesebb képjavítást tesz lehetővé.

Ebben a szakaszban bemutatunk egy sor különböző AI-alapú képjavító algoritmust, és bemutatjuk azok eredményeit. Ezek az algoritmusok kiterjednek a zajszűrésre, az élesség növelésre, a képek színminőségének javítására és még sok másra. Az AI-alapú képjavítás forradalmasította a képfeldolgozás területét, és széles körben alkalmazzák az orvosi képfeldolgozástól kezdve a fotómanipulációig, és számos más területen. Az itt bemutatott algoritmusok és eredmények segítenek megérteni ezen eljárások sokféleségét és hatékonyságukat a képek javítása terén. [16]

2.3.1. ArbSR

Az ArbSR (Arbitrary Super-Resolution) egy olyan mesterséges intelligencia alapú képjavító algoritmus, amelyet a "szuperfelbontás" (super-resolution) feladatokra terveztek. A szuperfelbontás azt a folyamatot jelenti, amikor egy alacsony felbontású képet (LR, Low-Resolution) magasabb felbontásúvá (HR, High-Resolution) alakítunk át úgy, hogy a képen lévő részleteket élesítjük és javítjuk. Az ArbSR különösen érdekes, mert lehetővé teszi a felhasználók számára, hogy a képfelbontást bármely választott értékre növeljék, nem korlátozva őket a megszokott szorzókhöz, mint például 2x, 4x stb.

Az ArbSR algoritmus alapja a mélytanulás és a neurális hálózatok használata, amelyek képesek megtanulni a képek tartalmát és összefüggéseit. A hálózatok képesek az alacsony felbontású képekhez tartozó magas felbontású párképek elemzésére, és megtanulják a kapcsolatot a két verzió között. Amikor egy alacsony felbontású képet adnak be az ArbSR algoritmusnak, az hálózat a tanult kapcsolatok alapján képes az új, magas felbontású változatot előállítani.

Az ArbSR algoritmus előnyei közé tartozik a sokoldalúsága, mivel a felhasználók szabadon választhatják a kívánt felbontást, és a képek minőségének hatékony javítása, amely lehetővé teszi a részletek kiemelését és az élesség növelését. Az ArbSR és hasonló szuperfelbontási algoritmusok egyre fontosabbak a fotómanipulációban, az orvosi képfeldolgozásban, a távérzékelésben és más területeken, ahol a képek minősége és részleteinek megjelenítése kulcsfontosságú. [17][18]

2.3.2. DiffBIR

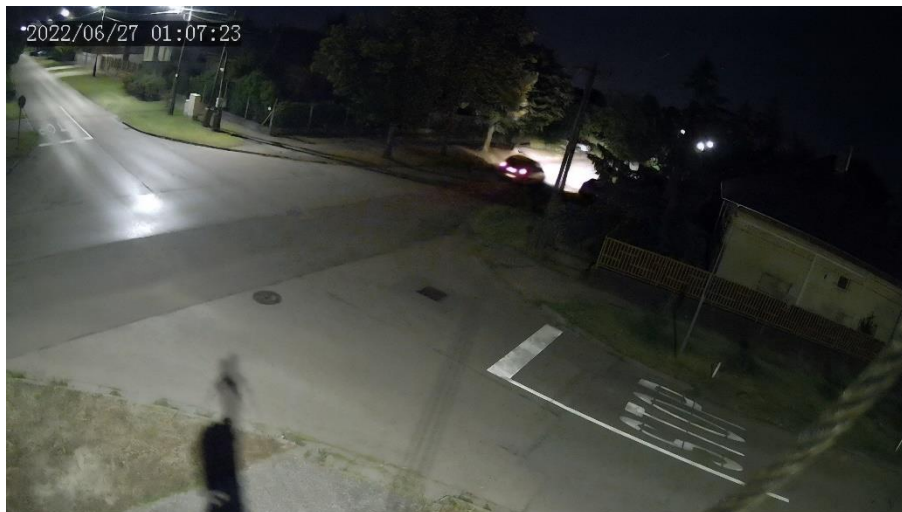
A DiffBIR (Differentiable Bayesian Image Reconstruction) egy olyan mesterséges intelligencia alapú képjavító algoritmus, amely a képjavítás területén a valószínűségi módszerekre és a mélytanulásra támaszkodik. A DiffBIR az egészségügyi képfeldolgozásban és az orvosi képfeldolgozásban talál széles körű alkalmazást, különösen a kép rekonstrukciós és zajszűrési feladatokban.

Az algoritmus kiindulópontja az, hogy az orvosi képek gyakran zajosak és hiányosak lehetnek, ami megnehezíti az orvosok számára a pontos diagnózist. A DiffBIR erre a problémára kínál megoldást azzal, hogy valószínűségi modellezést alkalmaz a képrekonstrukcióra. A modellezés során a DiffBIR becslést készít az eredeti kép reális eloszlásáról és a zaj jellemzőiről.

A DiffBIR alapjául szolgáló mechanizmus az ún. "variational autoencoder" (VAE) hálózatok és a Bayesi feltételes modell. Ezek a mechanizmusok lehetővé teszik a képek valószínűségi reprezentációjának létrehozását és a képek zajos változatának megbecsülését. Az algoritmus az elvárás-maximalizáció (expectation-maximization, EM) algoritmus alapján működik, amely a rekonstrukciós feladatot differenciálhatóvá teszi.

Az DiffBIR rendkívül hasznos lehet az orvosi képek újraépítésében és javításában. Például, a zajszűrés során a DiffBIR képes meghatározni a képeken lévő zaj jellegzetességeit és eltávolítani azt, anélkül, hogy elveszítené a fontos részleteket. Az algoritmus javíthatja az orvosi képek minőségét és pontosabb diagnózisokat eredményezhet.

A DiffBIR és ehhez hasonló valószínűségi alapú képjavító algoritmusok egyre növekvő figyelmet kapnak az egészségügyi képfeldolgozásban és más alkalmazási területeken, mivel segítenek javítani a képek minőségét és informativitását, ami kritikus fontosságú lehet a betegségek korai felismerésében és az orvosi vizsgálatok hatékonyságában. [19]



15. ábra: DiffBIR algoritmus

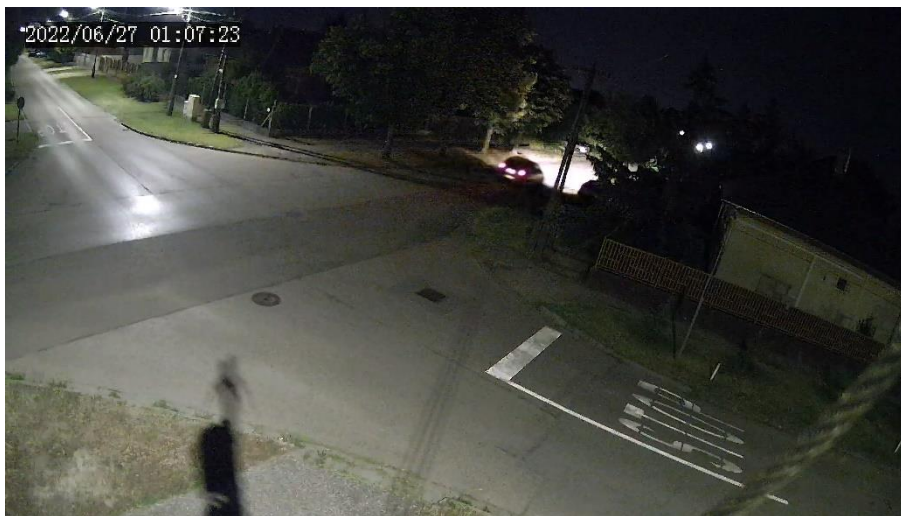
2.3.3. ESRGAN

Az ESRGAN (Enhanced Super-Resolution Generative Adversarial Network) egy olyan mélytanuláson alapuló képjavító algoritmus, amely a szuperfelbontási (super-resolution) feladatokra összpontosít. A szuperfelbontás során az alacsony felbontású képeket (LR, Low-Resolution) magas felbontásúvá (HR, High-Resolution) alakítjuk át úgy, hogy a képek részleteit javítjuk és élessé tesszük.

Az ESRGAN egy úgynevezett generatív adversariális hálózat (GAN) alkalmazása, amely két fő részből áll: egy generatív hálózathoz (Generative Network) és egy diszkriminatív hálózathoz (Discriminative Network). A generatív hálózat célja, hogy megtanulja a kapcsolatot az alacsony felbontású és magas felbontású képek között, és képes legyen előállítani a magas felbontású változatot az alacsony felbontású bemenet alapján. A diszkriminatív hálózat feladata pedig megítélni, hogy az előállított képek mennyire hasonlítanak az eredeti magas felbontású képekhez.

Az ESRGAN különösen hatékony az élesítési feladatokban, és képes kiemelni a képek részleteit és javítani a képek élességét. Az algoritmus eredményeképpen a képek sokkal részletesebbek és valóságosabbak lehetnek. Az ESRGAN a szuperfelbontás területén jelentős áttörést hozott, és hasznos eszköz lehet a fotómanipulációban, a filmes és játékgrafikában, valamint az orvosi képfeldolgozásban.

Az ESRGAN példája annak, hogy a mélytanulás és a GAN-technikák hogyan járulnak hozzá a képek minőségének és részleteinek javításához, ami számos alkalmazási területen előnyöket hozhat, ahol a vizuális minőség és a részletek kulcsfontosságúak. [20]



16. ábra: ESRGAN algoritmus

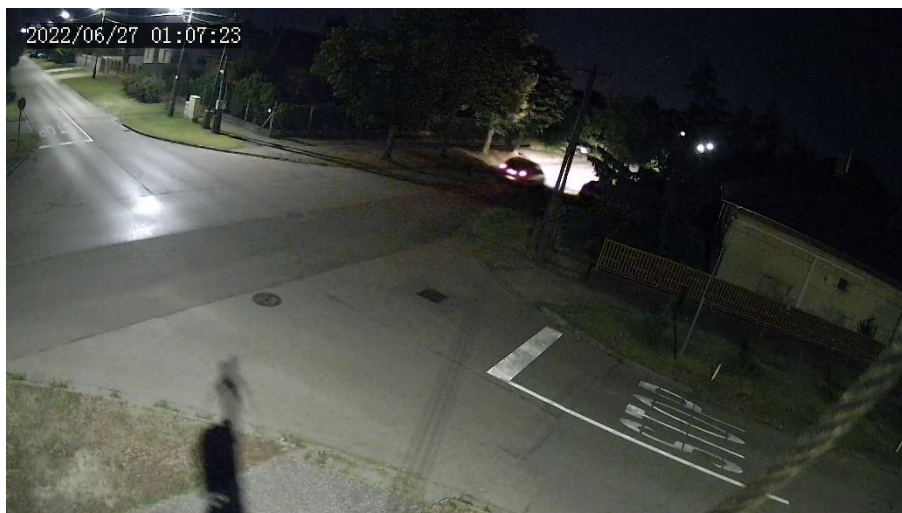
2.3.4. HAT

Az HAT (Holistic Attention Transformer) egy olyan mélytanuláson alapuló képjavító algoritmus, amely kiterjeszti a figyelmet az egész képre, hogy javítsa a képek minőségét és részleteit. Az algoritmus a figyelmet és a transzformátor architektúrát használja fel az élesítés és a részletek kiemelése érdekében.

A HAT alapja az Attention mechanizmus, amely lehetővé teszi a hálózat számára, hogy megtanulja, mely részletekre kell nagyobb hangsúlyt fektetni a képeken. Ezáltal az algoritmus képes az élesítésre és a képek részleteinek kiemelésére azokon a területeken, ahol a legnagyobb javulás szükséges.

A HAT-nek számos alkalmazási területe van, például a fotómanipuláció, a játékfejlesztés és a filmes effektek. Az algoritmus lehetővé teszi a képek részleteinek javítását és élesítését, ami vizuálisan vonzóbb és realisztikusabb eredményekhez vezethet.

A HAT egy példa arra, hogyan alkalmazhatók az Attention és a mélytanulási technikák a képfeldolgozásban a képek minőségének és részleteinek javítására, és hozzájárulhat a vizuális tartalmak fejlesztéséhez és javításához számos alkalmazási területen. [21]



17. ábra: HAT algoritmus

2.3.5. HCFlow-SR

A HCFlow-SR egy olyan mélytanuláson alapuló képjavító algoritmus, amely a szuperfelbontási (super-resolution) feladatokra összpontosít. A szuperfelbontás során az alacsony felbontású képek (LR, Low-Resolution) átalakításra kerülnek magas felbontásúvá (HR, High-Resolution), ahol a képek részleteinek és minőségének javítása a cél. [22]

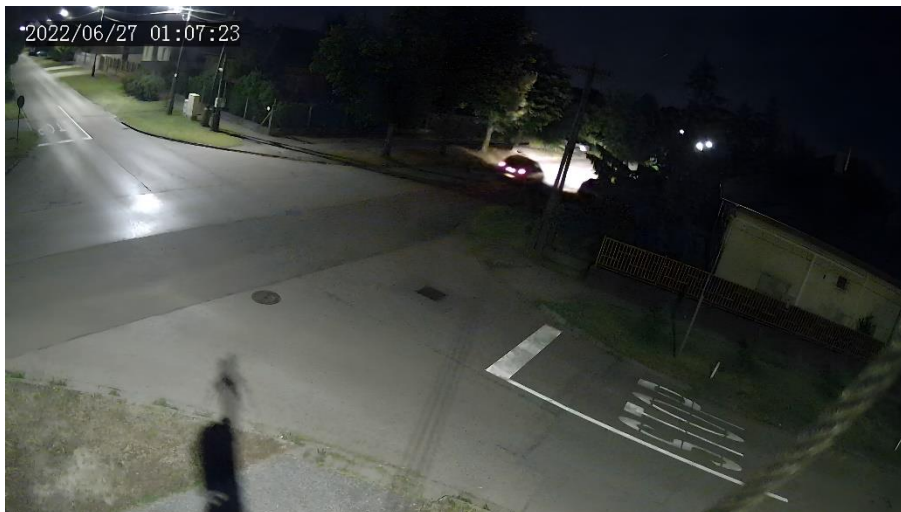
2.3.6. Real-ESRGAN

A Real-ESRGAN (Real-Enhanced Super-Resolution Generative Adversarial Network) szintén egy fejlett szuperfelbontási (super-resolution) algoritmus. Ez az algoritmus a generatív adversariális hálózatok (GAN) architektúráját alkalmazza a feladatra, ami lehetővé teszi a valósághűbb és részletesebb magas felbontású képek előállítását alacsony felbontású bemenetek alapján.

A Real-ESRGAN alapja az, hogy megtanulja a kapcsolatot az alacsony felbontású (LR) és a magas felbontású (HR) képek között. Az algoritmus egy generatív hálózattal rendelkezik, amely megpróbálja előállítani a magas felbontású változatot az alacsony felbontású képek alapján. Egy diszkriminatív hálózat is része az algoritmusnak, amelynek feladata megítélni, hogy az előállított képek mennyire hasonlítanak az eredeti magas felbontású képekhez.

A Real-ESRGAN kiemelkedő az élesség növelésében és a képek részleteinek javításában. Az algoritmus képes valósághűbb és részletesebb képek előállítására, amelyek vizuálisan vonzóbbak és minőségibbek. Az algoritmust a fotómanipulációban, a játékfejlesztésben, a filmes és televíziós produkciónál, valamint az orvosi képfeldolgozásnál alkalmazzák.

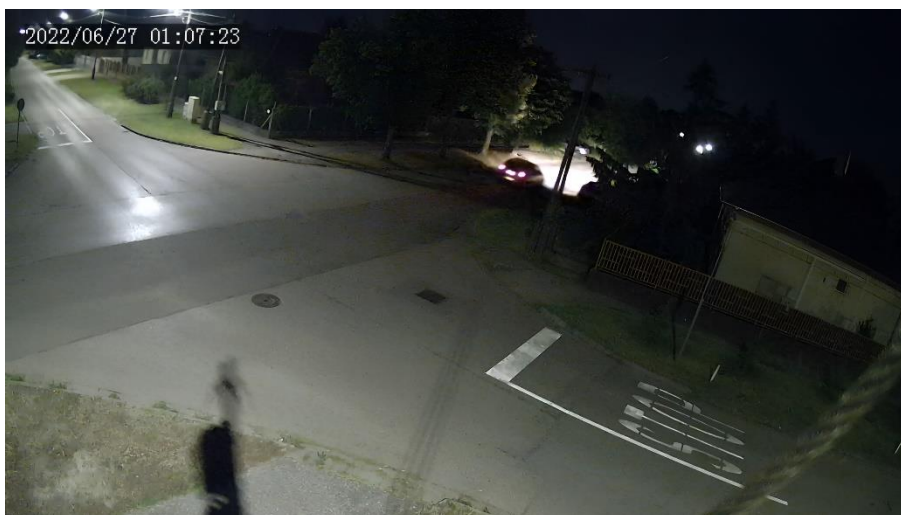
A Real-ESRGAN és ehhez hasonló GAN-alapú szuperfelbontási algoritmusok kiemelkedőek a képek minőségének javításában és a részletek kiemelésében, ami számos alkalmazási területen előnyt hozhat, ahol a vizuális minőség és a részletek kritikus fontosságúak. [23]



18. ábra: RealESRGAN algoritmus

2.3.7. ruDALL-E

A ruDALL-E algoritmus a Real-ESRGAN algoritmust alkalmazza, az elnevezése pedig a készítőkre utal. A Real-ESRGAN algoritmust fejlesztették tovább, amely által gyorsabbá és precízebbé vált.



19. ábra: ruDALL-E algoritmus

2.3.8. SRResCGAN

A SRResCGAN (Super-Resolution Residual Convolutional Generative Adversarial Network) egy intelligens képfelbontás növelésre szolgáló modell, amelynek hivatalos PyTorch implementációja a Deep Generative Adversarial Residual Convolutional Networks for Real-World Super-Resolution című tanulmányban található. A modell célja a valóságban

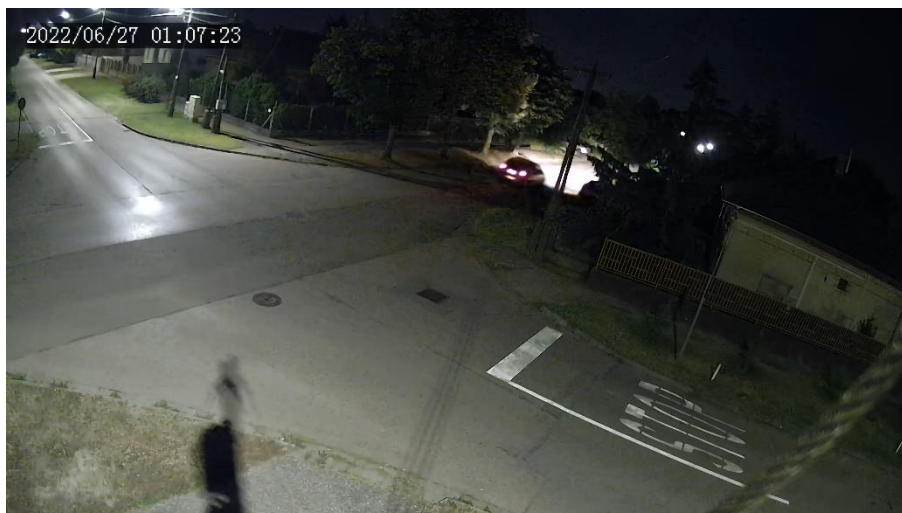
előforduló képfelbontás problémák kezelése, amelyek az alacsony felbontású (LR) képek és a magas felbontású (HR) képek közötti összefüggést zavarhatják. A jelenlegi mélytanulás alapú egyképes szuperfelbontás (SISR) módszerek általában feltételezik, hogy az LR képek a HR képek bicubikusan le-felbontott változatai. Azonban a valóságban ilyen lebontási folyamatok nem mindig állnak rendelkezésre, és az olyan tényezők, mint a szenzor zaj, a stochasztikus zaj, a tömörítési artefaktumok és a lehetséges különbségek az ábrázolás folyamata és a kamera eszköz között jelentősen rontják a jelenlegi SISR módszerek teljesítményét. Az SRResCGAN egy mély konvolúciós generatív adversariális hálózatot használ a probléma megoldására, amely valós világbeli korlátozásokat is figyelembe vesz azáltal, hogy adversariális módon tanítja a modellt pixel-szintű felügyelettel a generált LR verziójából a HR tartományban. Az előterjesztett hálózat kihasználja a maradékokkal való tanulást azzal, hogy minimalizálja az energián alapuló célállapot függvényt erőteljes kép regularizációs és konvex optimalizációs technikákkal. A módszert kvantitatív és minőségi kísérletekben mutatják be, amelyek robosztusan általánosítanak valós bemenetekre és könnyen alkalmazhatók más le-felbontási operátorokkal és mobil/beágyazott eszközökkel. [24]

2.3.9. Swin2SR

Swin2SR egy képfelbontási (super-resolution) algoritmus, amely mély tanuláson alapul. Az algoritmus célja az alacsony felbontású képek (LR, Low-Resolution) átalakítása magas felbontásúvá (HR, High-Resolution), hogy a képek részleteit javítsa és élessé tegye. Swin2SR különösen hatékony az élesség növelésében és a képek minőségének javításában.

Az eljárás alapja a Swin Transformer nevű architektúra, amely egy transzformátor-alapú mély tanulási hálózatot alkalmaz a képek felbontásának növelésére. A Swin2SR az alacsony felbontású bemeneteket használja, hogy magas felbontású képeket generáljon, és megtanulja a kapcsolatot az alacsony és magas felbontású képek között.

Az algoritmus különösen hasznos lehet a fényképek, a grafikai tervezés és más vizuális tartalmak feljavításában. A Swin2SR elősegíti a képek részleteinek és minőségének javítását, és valósághűbb eredményeket eredményezhet. Az ilyen algoritmusokat a fotómanipulációban, a játékfejlesztésben, a filmes és televíziós produkciókban, valamint az orvosi képfeldolgozásban használhatják. [25][26]



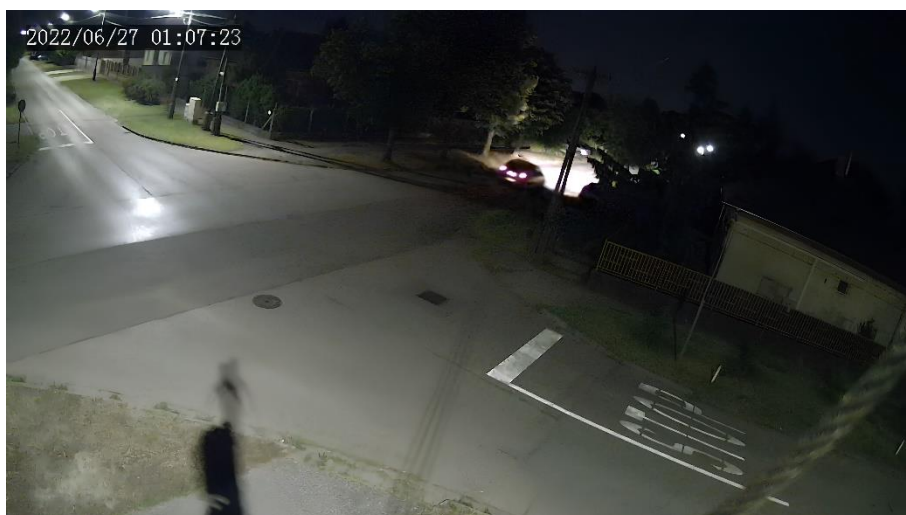
20. ábra: Swin2SR algoritmus

2.3.10. SwinIR

SwinIR egy mély tanuláson alapuló képfeljavító algoritmus, amely a képek minőségét és részleteit célozza meg. A név "SwinIR" valószínűleg az "Image Restoration" (képfeljavítás) rövidítése és a "Swin Transformer" architektúra nevének összekapcsolásából származik.

Az algoritmus lényege, hogy megtanulja a kapcsolatot az alacsony felbontású (LR, Low-Resolution) képek és a magas felbontású (HR, High-Resolution) képek között. A mély tanulási hálózatok segítségével az algoritmus képes az alacsony felbontású bemenetek alapján előállítani a magas felbontású változatokat.

A SwinIR alkalmazása lehetővé teszi a fényképek, képek és grafikák részleteinek javítását, ami fontos lehet vizuális tartalmak fejlesztése és javítása szempontjából. Az ilyen algoritmusokat széles körben alkalmazzák fotómanipulációban, játékfejlesztésben, filmes és televíziós termelésben, orvosi képfeldolgozásban, valamint más olyan alkalmazási területeken, ahol a képek minőségének és részleteinek jelentősége van. [27][28]



21. ábra: SwinIR algoritmus

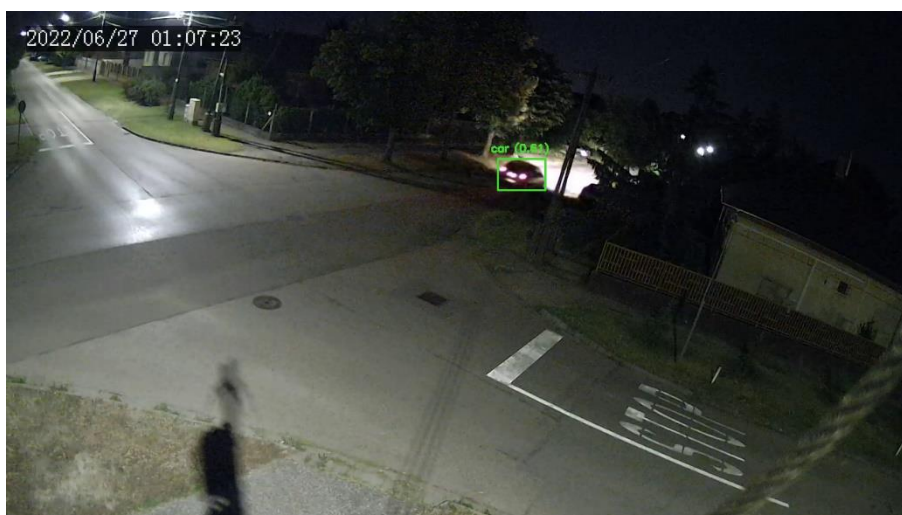
3. Képjavitó algoritmusok eredményei

A modern képfelismerés és objektumfelismerés területén jelentős előrelépések történtek az elmúlt években az mesterséges intelligencia (AI) és mély tanulás területén. Az objektumfelismerés kiemelkedő fontosságú számos alkalmazási területen, beleértve az önvezető járművek, a térfigyelés, a biztonságtechnika és az orvosi képfeldolgozás területét. Az objektumfelismerés terén a hagyományos algoritmusok mellett az AI alapú megközelítések is egyre inkább elterjedtek, és ígéretes eredményeket értek el.

Ebben a munkában a YOLOv8x modellt alkalmaztam az objektumok felismerésére, amely az egyik vezető AI alapú technológia az objektumfelismerés területén. A modell erőteljes képességekkel rendelkezik az objektumok detektálásában és kategorizálásában, valamint felhasználása számos gyakorlati alkalmazásban rendkívül fontos.

A munkám során összehasonlítottam a hagyományos algoritmusok és az AI alapú megközelítések eredményeit. A hagyományos algoritmusok közé tartoznak olyan módszerek, mint a kaskád osztályozók és OpenCV alapú technikák. Ezen algoritmusok teljesítményét metrikák, például pontosság, érzékenység és specifikusságuk alapján értékelem. Az AI alapú megközelítések esetében különös hangsúlyt fektettem a YOLOv8x modellre, és bemutatom annak eredményeit az objektumfelismerés terén.

Az 1. ábrán látható képen a YOLOv8x modell 61,41%-os valószínűséggel ismerte fel a gépjárművet, és helyesen autónak detektálta azt.



22. ábra: Alapkép eredménye

3.1. Hagományos algoritmusok

Az 2.2.-2.3. fejezetekben elvégzett kísérletek során számos képfelbontás és képfeldolgozás területén alkalmazott algoritmust értékeltem, hogy meghatározhassam,

melyik nyújtja a legjobb teljesítményt a gépjárművek felismerése során. A teszteket a YOLOv8x (You Only Look Once) modell segítségével végeztem, amely egy vezető gépi tanulás alapú technológia az objektumfelismerés területén.

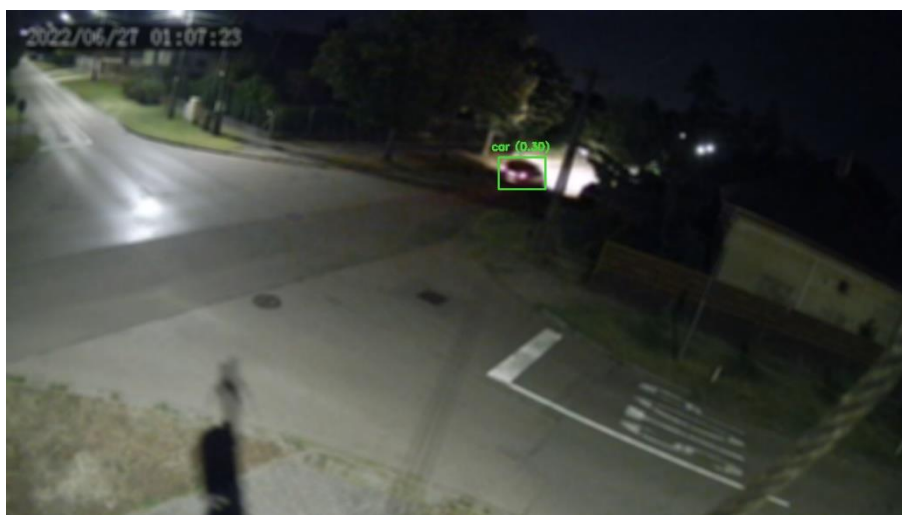
Az eredmények összefoglalója a következő:

1. táblázat: Hagyományos algoritmusok eredményei

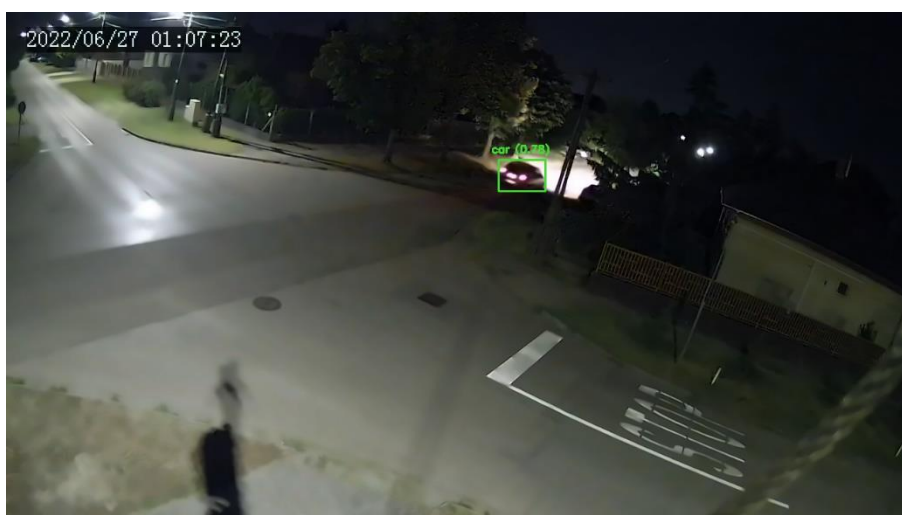
	Bilateral filter	Bilinear filter	Brightness I.	Brightness II.	Kontrasztolás	Decimation	Gamma
Autók száma	1	1	1	1	2	1	1
Pontosság	65,74%	62,3%	65,50%	65,10%	67,90%	62,30%	56,14%
Pontosság	-	-	-	-	62,18%	-	-
Hibák száma	1	0	1	1	1	0	0
	Gaussian kernel 3	Gaussian kernel 5	Gaussian kernel 9	Interpoláció	Szomszédos interpoláció	NLM	
Autók száma	1	1	1	1	1	1	
Pontosság	56,75%	53,49%	30,32%	61,41%	61,41%	77,98%	
Pontosság	-	-	-	-	-	-	
Hibák száma	0	0	0	0	0	0	

A fenti táblázat alapján az NLM (Non-Local Means) algoritmus nyújtotta a legmagasabb pontosságot a gépjárművek felismerése során. Ezzel szemben a Gaussian Filter 9-es méretű kernellel a legrosszabb teljesítményt produkálta. Az eredmények a korábban felállított hipotéziseimet támasztják alá. A különféle algoritmusok előnyeire és hátrányaira a 2. fejezetben tértem ki.

Ezenkívül az eredmények értékelése során figyelembe veszem az algoritmusok számítási erőforrásigényét, és elemzésüket kiterjesztem az objektumfelismerés gyakorlati alkalmazásaira.



23. ábra: Gauss-szűrő eredménye



24. ábra: NLM algoritmus eredménye

3.2. Mesterséges intelligencia alapú képjavító algoritmusok eredménye

A 2.3. pontban bemutatott AI alapú algoritmusokat is vizsgáltam az objektumfelismerés és képfeljavítás területén, és az eredmények bemutatása céljából összeállítottam egy összehasonlító táblázatot.

2. táblázat: AI alapú algoritmusok eredményei

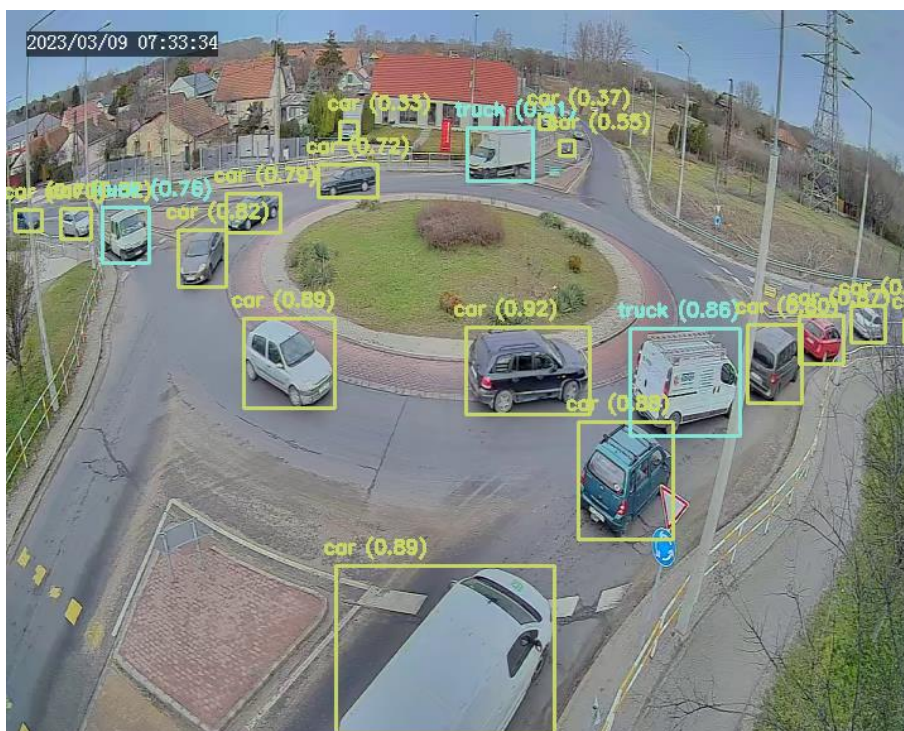
	DiffBIR	ESRGAN	HAT	Real-ESRGAN	ruDALL-E	Swin2SR	SwinIR
Autók száma	1	2	1	1	1	1	1
Pontosság	72,69%	66,26%	66,58%	74,22%	71,75%	74,12%	75,89%
Pontosság	-	25,48%	-	-	-	-	-
Hibák száma	1	1	1	1	0	2	0

A 2. táblázatban látható eredmények alapján az ESRGAN lett a leggyengébb, bár az 2 autót is felismert, de a 2. autót csak 25,48%-os biztonsággal ismerte fel. A SwinIR lett a legjobb algoritmus az eredmények alapján, mivel ez 75,89%-os pontossággal ismerte fel a számunkra releváns járművet.

Az ESRGAN által felismert második jármű egy kocsibejárón áll, és csak a jobb első sárvédő alapján állapította meg, hogy az valószínűleg egy autó lehet. Számomra ez az információ azonban nem lényeges, mivel csak a közúton közlekedő járműveket szeretném detektálni.

3.3. A hagyományos és AI alapú algoritmusok kombinálása

A kutatás során a hagyományos képfeldolgozó algoritmusok közül az NLM-t választottam, mivel ez bizonyult a legjobb eredményt elérőnek az eddigi tesztek során. Az AI alapú algoritmusok közül pedig a SwinIR-t választottam, mivel ennek a modellnek sikerült a legkiválóbb eredményeket produkálnia az objektumfelismerés tekintetében. A két különböző típusú algoritmus alkalmazása lehetővé teszi a hagyományos és a mesterséges intelligencia alapú megközelítések összehasonlítását, hogy átfogóbb és hatékonyabb eredményeket érjünk el a kutatási területen.



25. ábra: Körforgalomban készült kép

(Forrás: saját kép)

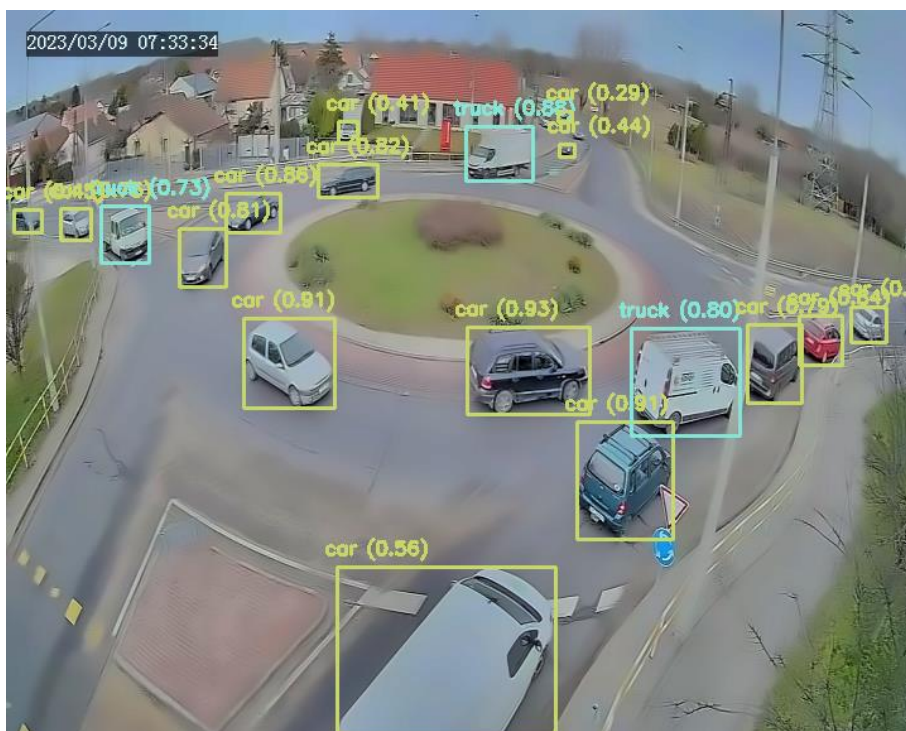


26. ábra: Forgalmas kereszteződés

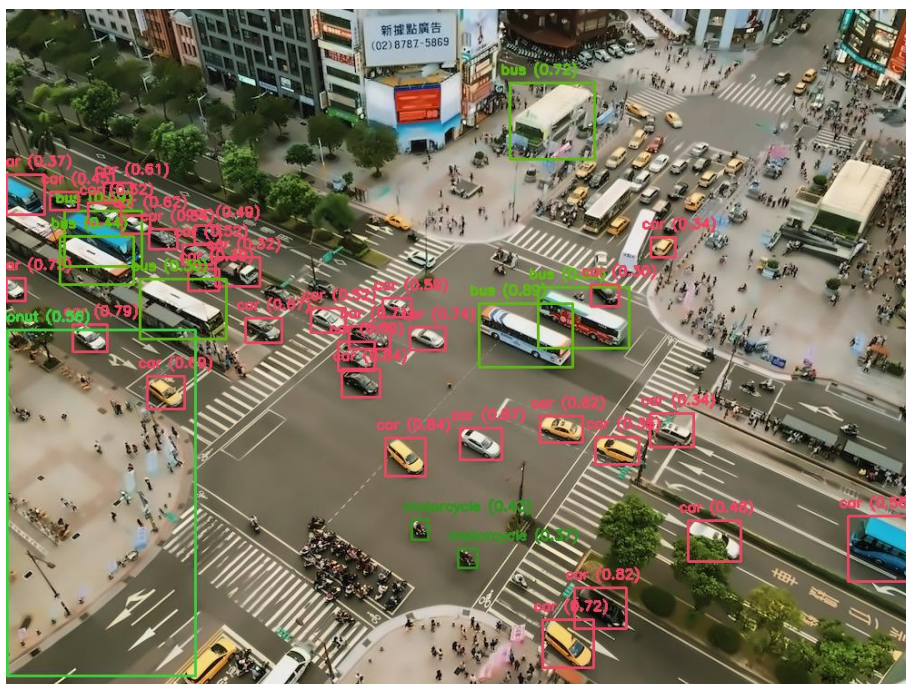
(Forrás: <https://images.pexels.com/photos/186537/pexels-photo-186537.jpeg>)

Először az alapképeken végeztem el az objektumdetektációt. A 25. ábrán látható, hogy a modell szinte az összes objektumot sikeresen felismerte. A 26. ábrán megfigyelhető, hogy a YOLOv8x modell nem tudta az összes objektumot felismerni, ennek az oka nagyon egyszerű, a gépjárművek olyan szögben látszódnak, amely már nem teljesen egyértelmű, ezért nem képes hatásosan felismerni azokat. Ez a tulajdonsága a Transfer Learning által javítható, és megszüntethető. A TL azt jelenti, hogy az előre tanított (pretrained) modellhez hozzáadjuk a saját mintáinkat, tovább tanítjuk azt.

A 27-28. ábrákon az eredeti képeken alkalmazott Non-Local Means (NLM) algoritmussal javított változatok láthatóak. A körforgalomban készült felvételen észlelhetők érdekes változások, ahol a járművek felismerési valószínűségeik változása megfigyelhető. A NLM-alapú javítás bizonyos járművek esetében csökkentette a felismerési valószínűséget, míg más járművek esetében javította azt. A nagy kereszteződés esetében a hibák száma kettőre csökkent, és a felismerési pontosságok szinte minden esetben növekedtek. Ez az eredmény azt sugallja, hogy a NLM algoritmus hatékonyan képes kijavítani és optimalizálni a képek objektumfelismerési pontosságát, különösen a nagyobb és komplexebb forgalmi környezetekben.



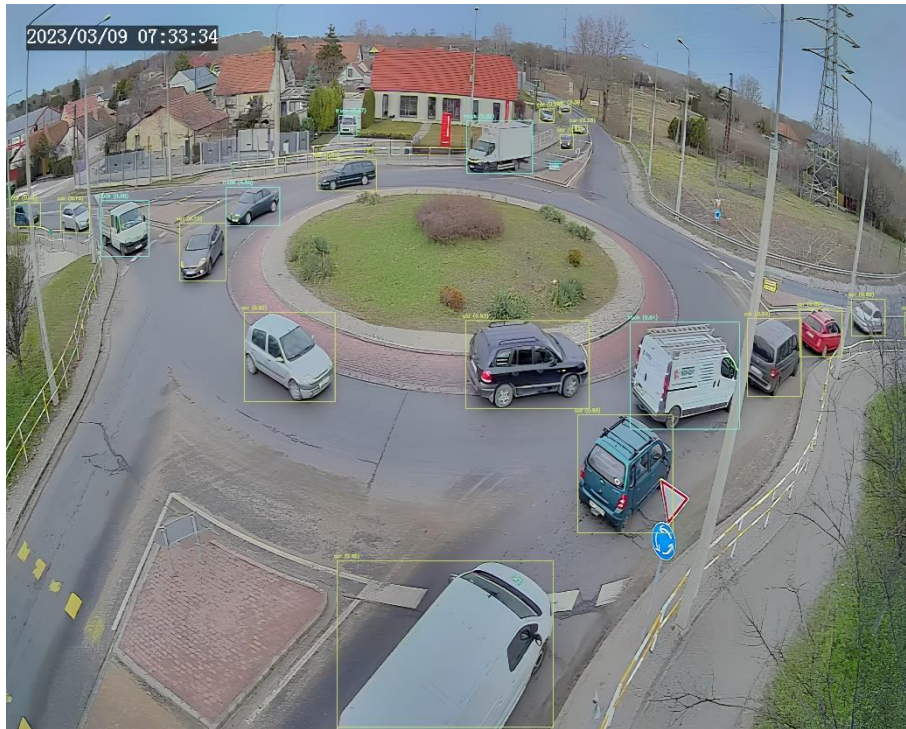
27. ábra: Körforgalmas kép NLM algoritmussal



28. ábra: Forgalmas kereszteződés NLM algoritmussal

A 29-30. ábrákon látható eredeti képeken alkalmazott SwinIR algoritmus által javított változatok érdekes eredményeket mutatnak. A nagy kereszteződés esetében három hiba figyelhető meg, azonban az eredmények minimálisan jobbak az NLM algoritmushoz képest. A körforgalomban készült felvételen pozitív változások tapasztalhatók, ahol a SwinIR algoritmus több objektumot ismert fel, és azok felismerési pontossága is szignifikánsan javult. Ezek az eredmények azt mutatják, hogy a SwinIR algoritmus hatékonyan képes

finomhangolni a képek objektumfelismerési képességeit, különösen az olyan forgalmi helyzetekben, ahol több és változatosabb objektumok vannak jelen.



29. ábra: Körforgalmas kép SwinIR algoritmussal



30. ábra: Forgalmas kereszteződés SwinIR algoritmussal



31. ábra: Körforgalmas kép NLM és SwinIR algoritmussal



32. ábra: Forgalmas kereszteződés NLM és SwinIR algoritmussal

A 31-32. ábrákon bemutatott eredeti képeken alkalmazott Non-Local Means (NLM), majd pedig a SwinIR algoritmusok által javított változatok érdekes eredményeket mutatnak. A kereszteződés esetében csak egy hiba figyelhető meg, ugyanakkor néhány félre detektálás is megfigyelhető. A körforgalom esetében az eredmények alapvetően megtartották az eredeti állapotot, a detektációk számát tekintve. Azonban mindkét esetben megfigyelhető, hogy a közeli objektumok felismerése során mindkét algoritmus magasabb pontossággal teljesít,

míg a távoli objektumoknál a felismerési pontosság csökken. Ez a jelenség felveti a távolság és objektumfelismerés közötti kihívásokat, különösen a távolabbi objektumok esetében.



33. ábra: Körforgalmas kép SwinIR és NLM algoritmussal



34. ábra: Forgalmás kereszteződés SwinIR és NLM algoritmussal

A 33-34. ábrákon a SwinIR, majd a Non-Local Means (NLM) algoritmusokkal javított változatok láthatóak. Az eredmények alapján minimális javulás figyelhető meg, de összességében a képek tartalmi elemei nagyrészt megtartották eredeti formájukat. A két algoritmus alkalmazásának sorrendje nem hozott szignifikáns különbséget az

objektumfelismerési pontosságokban vagy a hibák számában. Ez a jelenség azt sugallja, hogy mindkét algoritmus hasonlóan teljesít, és az alkalmazásuk sorrendje kevés hatással van a végeredményre a vizsgált képek esetében.

3. táblázat: YOLOv8x modell által felismert objektumok száma különböző képjavító algoritmus esetén

Algoritmus	Szemponatok (db)	Körforgalom	Kereszteződés
<i>Alap kép</i>	<i>Autók</i>	16	28
	<i>Kamionok/buszok</i>	3	5
	<i>Motorosok</i>	0	2
	<i>Egyebek</i>	0	1
	<i>Félredetektálás</i>	1	2
	<i>Hibás detektálás</i>	0	3
<i>NLM algoritmus</i>	<i>Autók</i>	15	31
	<i>Kamionok/buszok</i>	3	6
	<i>Motorosok</i>	0	2
	<i>Egyebek</i>	0	1
	<i>Félredetektálás</i>	1	3
	<i>Hibás detektálás</i>	0	1
<i>SwinIR algoritmus</i>	<i>Autók</i>	16	34
	<i>Kamionok/buszok</i>	4	9
	<i>Motorosok</i>	0	2
	<i>Egyebek</i>	0	4
	<i>Félredetektálás</i>	2	2
	<i>Hibás detektálás</i>	0	4
<i>NLM + SwinIR algoritmus</i>	<i>Autók</i>	15	37
	<i>Kamionok/buszok</i>	3	7
	<i>Motorosok</i>	0	2
	<i>Egyebek</i>	0	0
	<i>Félredetektálás</i>	0	2
	<i>Hibás detektálás</i>	0	1
<i>SwinIR + NLM algoritmus</i>	<i>Autók</i>	14	36
	<i>Kamionok/buszok</i>	4	8
	<i>Motorosok</i>	0	2
	<i>Egyebek</i>	0	6
	<i>Félredetektálás</i>	0	5
	<i>Hibás detektálás</i>	0	2

Összefoglalás

Az előzetes hipotéziseim alapján a gépi tanulás alapú objektumfelismerő algoritmusok eredményesebbek lehetnek a gépjárművek azonosításában, különösen változatos fényviszonyok és forgalmi helyzetek mellett. Azonban a végzett kísérletek és azok eredményei azt mutatták, hogy ez a hipotézis nem mindig igaz. A gépi tanulás alapú algoritmusok változatosan teljesítettek, és meglepő volt látni, hogy az ESRGAN algoritmus esetén, csak korlátozott sikerrel azonosított gépjárműveket a YOLOv8x modell, néhány esetben alacsony valószínűséggel és hibás detekcióval.

Ugyanakkor megnyugtató eredményeket tapasztaltunk a SwinIR algoritmus használatával, ami kiemelkedő teljesítményt mutatott a gépjárművek azonosításában. A 75,89% -os bizonyossággal és hibás detekciók nélkül elért eredmények azt sugallják, hogy a megfelelően kiválasztott és optimalizált gépi tanulás alapú algoritmusok hatékonyak lehetnek ezen a területen.

Fontos megjegyezni, hogy az algoritmusok teljesítménye nagyban függ az alkalmazási területtől, a kameraoptikától és a használt adathalmaztól. Az eredmények azt mutatják, hogy nincs univerzális "jobb" vagy "rosszabb" algoritmus az objektumfelismerésben, és a választásnak az adott környezet és adatok alapján kell történnie.

Ezen felül, az éjszakai és nappali felvételek esetében tapasztalt eltérések érdekesek voltak. A NLM algoritmus hatékonyabbnak bizonyult az éjszakai képek esetében, míg a SwinIR a nappaliaknál teljesített jobban. Azonban fontos megjegyezni, hogy a két algoritmus kombinált alkalmazása csak a közeli objektumokon hozott javulást, míg a távoli objektumoknál a teljesítmény romlott.

Az összességében szerzett tapasztalatok rámutatnak arra, hogy az objektumfelismerés terén nincsenek univerzális megoldások, és további kutatásokra és fejlesztésekre van szükség ahhoz, hogy még hatékonyabb és megbízhatóbb algoritmusokat hozzunk létre a gyakorlati alkalmazások számára.

Ábrajegyzék

1. ábra: Alapkép.....	8
2. ábra: Medián-szűrő alkalmazása	9
3. ábra: Gauss-szűrő alkalmazása 5-ös méretű kernellel	9
4. ábra: Gauss-szűrő alkalmazása 9-es méretű kernellel	10
5. ábra: Bilaterális szűrő alkalmazása	10
6. ábra: Non-local means szűrő alkalmazása	11
7. ábra: Fényerősség növelése I.	12
8. ábra: Fényerősség növelése II.....	12
9. ábra: Gammakorrektúrák algoritmus	13
10. ábra: Kontrasztosztás algoritmus.....	13
11. ábra: Interpoláció kétszeres nagyításban	15
12. ábra: Decimation algoritmus félszeres kicsinyítés	15
13. ábra: Bilineáris interpoláció kétszeres nagyításban	16
14. ábra: Közelítő interpoláció kétszeresen nagyítva	16
15. ábra: DiffBIR algoritmus.....	20
16. ábra: ESRGAN algoritmus	21
17. ábra: HAT algoritmus.....	22
18. ábra: RealESRGAN algoritmus.....	23
19. ábra: ruDALL-E algoritmus	23
20. ábra: Swin2SR algoritmus.....	25
21. ábra: SwinIR algoritmus.....	26
22. ábra: Alapkép eredménye	27
23. ábra: Gauss-szűrő eredménye	29
24. ábra: NLM algoritmus eredménye.....	29
25. ábra: Körforgalomban készült kép	30
26. ábra: Forgalmas kereszteződés	31
27. ábra: Körforgalmas kép NLM algoritmussal.....	32
28. ábra: Forgalmas kereszteződés NLM algoritmussal.....	32
29. ábra: Körforgalmas kép SwinIR algoritmussal	33
30. ábra: Forgalmas kereszteződés SwinIR algoritmussal	33
31. ábra: Körforgalmas kép NLM és SwinIR algoritmussal	34
32. ábra: Forgalmas kereszteződés NLM és SwinIR algoritmussal	34
33. ábra: Körforgalmas kép SwinIR és NLM algoritmussal	35
34. ábra: Forgalmas kereszteződés SwinIR és NLM algoritmussal	35

Irodalomjegyzék

- [1] Horváth Cz. J., Lévai J., Molnár Gy.: Oktatás és technológia, Typotex, 2014
- [2] N. Khandelwal: Image Processing in Python: Algorithms, Tools, and Methods You Should Know, Elektronikus forrás: <https://neptune.ai/blog/image-processing-python>, Megtekintés dátuma: 2023. 09. 25.
- [3] Pre-Scient Technologies: Image Processing Algorithms, Elektronikus forrás: <https://www.pre-scient.com/knowledge-center/image-processing/image-processing-algorithms.html>, Megtekintés dátuma: 2023. 09. 25.
- [4] Maklár Z.: Pixelszám, képek méretezése, tárhely méret, felbontás, Elektronikus forrás: https://www.fenykepes.hu/html/z10_kepmeretezes.html, Megtekintés dátuma: 2023. 09. 30.
- [5] Tanács A.: Képfeldolgozás Python + OpenCV környezetben, Elektronikus forrás: https://www.inf.u-szeged.hu/~tanacs/pyocv/zajterhels_s_zajszrs_kptrben.html, Megtekintés dátuma: 2023. 09. 30.
- [6] Csetverikov D.: Digitális képelemzés alapvető algoritmusai, Elektronikus forrás: https://www.inf.elte.hu/dstore/document/297/Csetverikov_jegyzet.pdf, Megtekintés dátuma: 2023. 10. 07.
- [7] R. Fisher, S. Perkins, A. Walker, E. Wolfart: HYPERMEDIA IMAGE PROCESSING REFERENCE – Median Filter, Elektronikus forrás: <https://homepages.inf.ed.ac.uk/rbf/HIPR2/median.htm>, Megtekintés dátuma: 2023. 10. 10.
- [8] R. Fisher, S. Perkins, A. Walker, E. Wolfart: HYPERMEDIA IMAGE PROCESSING REFERENCE – Gaussian Smoothing, Elektronikus forrás: <https://homepages.inf.ed.ac.uk/rbf/HIPR2/gsmooth.htm>, Megtekintés dátuma: 2023. 10. 10.
- [9] R. Fisher, S. Perkins, A. Walker, E. Wolfart: HYPERMEDIA IMAGE PROCESSING REFERENCE – Mean Filter, Elektronikus forrás: <https://homepages.inf.ed.ac.uk/rbf/HIPR2/mean.htm>, Megtekintés dátuma: 2023. 10. 10.
- [10] Berke D., Major K., Ocskai Zs., Enyedi A.: Digitális képkalkáló algoritmusok összehasonlító elemzése képszerkezet és entrópia alapján, Elektronikus forrás: https://www.researchgate.net/profile/Jozsef-Berke-2/publication/301845822_Digitalis_kepalkoto_algoritmusok_osszehasonlito_elemezese_kepszerkezet_es_entropia_alapjan/links/572a0af108ae057b0a078669/Digitalis-kepalkoto-algoritmusok-oesszehasonlito-elemzese-kepszerkezet-es-entropia-alapjan.pdf, Megtekintés dátuma: 2023. 10. 16.
- [11] Kovács Gy., Fazekas A.: Automatikus gamma korrekció, Elektronikus forrás: https://www.inf.u-szeged.hu/projectdirs/kepaf2011/pdfs/S04_04.pdf, Megtekintés dátuma: 2023. 10. 17.
- [12] D. Etohells: Understanding gamma correction, Elektronikus forrás: <https://www.cambridgeincolour.com/tutorials/gamma-correction.htm>, Megtekintés dátuma: 2023. 10. 17.

- [13] Pixinfo: Interpolációs eljárások, Elektronikus forrás: <https://pixinfo.com/cikkek/interpolacios-eljarasok/>, Megtekintés dátuma: 2023. 10. 18.
- [14] Enyedi A., Kozma-Bognár V., Berke J.: Távérzékelési célú képalkotó algoritmusok összehasonlítása tartalom és szerkezet alapján, Elektronikus forrás: <https://www.rsgis.hu/RS&GIS-2016-6.pdf>, Megtekintés dátuma: 2023. 10. 17.
- [15] Szemenyei M.: Számítógépes Látórendszerek, Elektronikus forrás: https://vik.wiki/images/1/15/Szl_jegyzet.pdf, Megtekintés dátuma: 2023. 10. 19.
- [16] M. Samir: Unveiling the Latest Algorithms in Image Processing: Unlocking the Power of Images, Elektronikus forrás: <https://www.linkedin.com/pulse/unveiling-latest-algorithms-image-processing-unlocking-md-samir>, Megtekintés dátuma: 2023. 10. 20.
- [17] X. Wang és mtsai: Deep Arbitrary-Scale Image Super-Resolution via Scale-Equivariance Pursuit, Elektronikus forrás: https://openaccess.thecvf.com/content/CVPR2023/papers/Wang_Deep_Arbitrary-Scale_Image_Super-Resolution_via_Scale-Equivariance_Pursuit_CVPR_2023_paper.pdf, Megtekintés dátuma: 2023. 10. 21.
- [18] The Learning and Vision Atelier: ArbSR, Elektronikus forrás: <https://github.com/The-Learning-And-Vision-Atelier-LAVA/ArbSR/blob/master/README.md>, Megtekintés dátuma: 2023. 10. 21.
- [19] X. Lin és mtsai: DiffBIR : Towards Blind Image Restoration with Generative Diffusion Prior, Elektronikus forrás: https://www.researchgate.net/publication/373487536_DiffBIR_Towards_Blind_Image_Restoration_with_Generative_Diffusion_Prior, Megtekintés dátuma: 2023. 10. 22.
- [20] X. Wang és mtsai: ESRGAN: Enhanced Super-Resolution Generative Adversarial Networks, Elektronikus forrás: https://openaccess.thecvf.com/content_ECCVW_2018/papers/11133/Wang_ESR_GAN_Enhanced_Super-Resolution_Generative_Adversarial_Networks_ECCVW_2018_paper.pdf, Megtekintés dátuma: 2023. 10. 22.
- [21] X. Chen és mtsai: HAT: Hybrid Attention Transformer for Image Restoration, Elektronikus forrás: https://www.researchgate.net/publication/373837869_HAT_Hybrid_Attention_Transformer_for_Image_Restoration, Megtekintés dátuma: 2023. 10. 23.
- [22] J. Liang és mtsai: HCFlow, Elektronikus forrás: <https://github.com/JingyunLiang/HCFlow/blob/main/README.md>, Megtekintés dátuma: 2023. 10. 23.
- [23] X. Wang és mtsai: Real-ESRGAN: Training Real-World Blind Super-Resolution with Pure Synthetic Data, Elektronikus forrás: <https://arxiv.org/pdf/2107.10833.pdf>, Megtekintés dátuma: 2023. 10. 23.
- [24] R. M. Umer, G. L. Foresti, C. Micheloni: Deep Generative Adversarial Residual Convolutional Networks for Real-World Super-Resolution, Elektronikus forrás: https://openaccess.thecvf.com/content_CVPRW_2020/papers/w31/Umer_Deep_Generative_Adversarial_Residual_Convolutional_Networks_for_Real-

- [World Super-Resolution CVPRW 2020 paper.pdf](#), Megtekintés dátuma: 2023. 10. 24.
- [25] MV-Lab: Swin2SR, Elektronikus forrás: <https://github.com/mv-lab/swin2sr/blob/main/README.md>, Megtekintés dátuma: 2023. 10. 24.
- [26] Z. Liu és mtsai: Swin2SR, Elektronikus forrás: https://huggingface.co/docs/transformers/v4.31.0/model_doc/swin2sr, Megtekintés dátuma: 2023. 10. 25.
- [27] J. Liang: SwinIR: Image Restoration Using Swin Transformer, Elektronikus forrás: https://openaccess.thecvf.com/content/ICCV2021W/AIM/papers/Liang_SwinIR_Image_Restoration_Using_Swin_Transformer_ICCVW_2021_paper.pdf, Megtekintés dátuma: 2023. 10. 26.
- [28] N.-L- Nguyen: A Brief Analysis of the SwinIR Image Super-Resolution, Elektronikus forrás: <https://www.ipol.im/pub/art/2022/430/article.pdf>, Megtekintés dátuma: 2023. 10. 27.