

# AI trendek a digitális képfeldolgozásban

## AI Trends in Digital Image Processing

**Dr. VARGA Attila Károly**  
egyetemi docens

Miskolci Egyetem  
Gépészmérnöki és Informatikai Kar  
Automatizálási és Infokommunikációs Intézet  
3515 Miskolc-Egyetemváros  
+36 46 565 111 / 19-46  
attila.varga@uni-miskolc.hu  
www.uni-miskolc.hu

### Abstract

*In digital image processing, artificial intelligence is increasingly used for image analysis, enhancement, pattern recognition, object recognition and classification. In contrast to traditional image processing, which is often based on rules and predefined algorithms, AI-based approaches rely on learning, adaptation and automatic decision making to recognise and manage image features. Core technologies include deep learning, neural networks and machine learning)-based algorithms. AI-based technology is already present in a growing number of fields and industries, significantly complementing classical approaches to image processing or replacing some of its steps or sub-processes with the power of machine intelligence.*

**Keywords:** digital image processing, artificial intelligence, pattern recognition, deep learning, machine learning

### Kivonat

*A digitális képfeldolgozás területén a mesterséges intelligenciát egyre gyakrabban használják képek elemzésére, javítására, mintafelismerésre, objektumok azonosítására és képosztályozásra. A hagyományos képfeldolgozással szemben, amely gyakran szabályok és előre meghatározott algoritmusok alapján működik, az AI-alapú megközelítés tanulásra, adaptációra és automatikus döntéshozatalra épít a képek jellemzőinek felismerésében és kezelésében. Az alapvető technológiák közé tartozik a mélytanulás, a neurális hálózatok és a gépi tanulás alapú algoritmusok. Az AI alapú technológia már egyre több területen, iparágban jelen van, a képfeldolgozás klasszikus megközelítéseit jelentősen kiegészítve vagy annak egyes lépéseit, részfolyamatát helyettesítve a gépi intelligencia erejével.*

**Kulcsszavak:** digitális képfeldolgozás, mesterséges intelligencia, mintafelismerés, mélytanulás, gépi tanulás

## 1. BEVEZETÉS

Az AI alapú digitális képfeldolgozás olyan technikákat és algoritmusokat alkalmaz, amelyek mesterséges intelligenciára, különösen gépi tanulásra épülnek. A mesterséges intelligencia alapú algoritmusok lehetővé teszik a képfeldolgozási folyamatok automatizálását, ami jelentősen felgyorsíthatja a munkafolyamatokat, különösen nagy mennyiségű adat esetén. A digitális képfeldolgozás alapvetően képek számítógépes elemzését és manipulálását jelenti, amely lehetőséget nyújt olyan műveletekre [1], mint:

- képjavítás,
- képszegmentálás (objektumok elkülönítése a képen),
- képosztályozás (pl. képen lévő objektum felismerése),
- képrekonstrukció,
- vagy képpalkotó rendszerek felbontásának növelése.

A mesterséges intelligencia képes arra, hogy a képeken felismerje az objektumokat, és kategóriákba sorolja őket, az AI modellek képesek megtanulni a képeken belüli mintázatokat, és ennek alapján egyre pontosabb eredményeket nyújtanak. Például, egy mély neurális hálózat megtanulhatja, hogy felismerje, mi van egy képen és kategorizálja azokat. A szegmentálás során az AI azonosítja és elkülöníti a képen belüli különböző régiókat. Ez rendkívül fontos például az orvosi képfeldolgozásban, ahol különböző szöveteket, szerveket vagy elváltozásokat kell elkülöníteni egy röntgen- vagy MRI-felvételen. Az AI képes hatékonyabb tömörítési technikák fejlesztésére, amelyek csökkentik a kép méretét, miközben megőrzik annak minőségét. Az automatikus tömörítési módszerek mesterséges intelligenciával képesek jobban optimalizálni a képi adatokat, mint a hagyományos módszerek. A generatív AI [2] alkalmazásával egy képet egy másik stílusban lehet újragenerálni, például egy festmény stílusában átalakíthatunk egy fotót, amely így egy adott művész által festett képhez fog hasonlítani. Az AI-alapú képfeldolgozás lehetővé teszi az objektumok felismerését és követését videókban vagy élő kameraképeken. Ez olyan területeken hasznos, mint például a biztonságtechnika (pl. arcfelismerés), az autonóm járművek vagy okos városok rendszerei.

## 2. AI ALAPÚ ALGORITMUSOK A KÉPFELDOLGOZÁSOKBAN

Az AI-alapú digitális képfeldolgozásban számos új megoldás azonosítható, amelyek az utóbbi évek technológiai fejlődésével párhuzamosan jelentek meg. Ezek a trendek a mesterséges intelligencia (AI), a gépi tanulás (ML) és a deep learning (mélytanulás) alkalmazásának köszönhetően formálják a képfeldolgozás területét. [3]

### 2.1. Mély neurális hálózatok széles körű alkalmazása

A konvolúciós neurális hálózatok (CNN) terjedése a képfeldolgozás terén a legjelentősebb trend. A deep learning módszerek rendkívül jól teljesítenek képfelismerésben, szegmentálásban, objektumfelismerésben és más képanalízis feladatokban. [4]

A konvolúciós neurális hálózat egy speciális mélytanulási modell, amelyet kifejezetten képek és más hasonló típusú, strukturált adatok feldolgozására terveztek. A CNN-ek azzal váltak népszerűvé, hogy hatékonyan tudják felismerni a képeken található mintázatokat, például éleket, textúrákat, formákat és objektumokat, így ideálisak képfeldolgozási feladatokra, mint például az objektumfelismerés, képosztályozás és képszegmentálás.

A CNN-ek alapvető működési elve a konvolúció, amely egy matematikai művelet, amely a kép lokális jellemzőit (pl. a pixelek eloszlását) dolgozza fel és felismeri a releváns mintákat. A konvolúciós neurális hálózatok a hagyományos neurális hálózatoktól eltérően olyan rétegekkel dolgoznak, amelyek speciális műveleteket végeznek a bemeneti képen. A CNN több egymásra épülő rétegből áll, amelyek a kép egyre összetettebb jellemzőit tanulják meg.

A CNN alapvető építőeleme a konvolúciós réteg, amely egy konvolúciós műveletet hajt végre a bemeneti képen. A konvolúciós művelet során egy szűrőt (vagy más néven kernel) "csúsztatnak" végig a képen, és minden pozíciónál egy kis régiót vesz figyelembe a képből. A szűrő célja, hogy bizonyos mintázatokat vagy jellemzőket, például éleket vagy sarkokat ismerjen fel. A szűrők mérete általában kisebb, mint a teljes kép (például 3x3 vagy 5x5 pixeles). A szűrő minden pozícióban egy pontértéket generál, és így hoz létre egy új képet (feature map), amely az adott szűrő által felismert jellemzőket tartalmazza.

Minden konvolúciós művelet után egy nemlinearitást alkalmaznak a kimenetre, amelyet ReLU aktivációs függvénynek nevezünk. A ReLU célja, hogy eltávolítsa a negatív értékeket, így segít a hálózatnak a nemlineáris kapcsolatok megtanulásában.

A pooling réteg feladata, hogy az információ sűrítésével csökkentse a konvolúciós réteg által előállított kép méretét. Ezáltal a hálózat hatékonyabban dolgozik, és kevésbé érzékeny a bemeneti kép apróbb változásaira. A leggyakoribb pooling technika a max pooling, amely a kép egy kis részén (például 2x2 vagy 3x3 ablakon) belül kiválasztja a legnagyobb értéket, és azt viszi tovább a következő rétegbe.

A CNN utolsó része általában néhány teljesen összekapcsolt rétegből áll, amely hasonló a hagyományos neurális hálózatokban található rétegekhez. Itt minden neuronnak kapcsolata van minden más neuronnal az előző rétegből. Ez a rész végzi el a tényleges osztályozási feladatot, például eldönti, hogy egy képen egy

macska, kutya vagy autó található-e. Az utolsó réteg a kimeneti réteg, amely gyakran egy softmax függvényt használ, hogy az osztályozási feladat esetén megadja a különböző osztályokhoz tartozó valószínűségeket. A CNN-ek a képek lokális jellemzőit dolgozzák fel, ami lehetővé teszi számukra, hogy kis mintázatokat (pl. éleket, textúrákat) könnyen felismerjenek. Ez sokkal hatékonyabbá teszi őket, mint a hagyományos neurális hálózatok, amelyek minden egyes pixelt külön kezelnek. A CNN-ek szűrői ugyanazokat a súlyokat használják az egész képre, ami csökkenti a paraméterek számát, így kevesebb adattal is képesek megtanulni a jellemzőket. A deep learning algoritmusok képesek automatikusan felismerni és kinyerni a releváns jellemzőket a képekből, ami forradalmasítja a képfeldolgozást az orvosi diagnosztikától az autonóm járművekig. [5][6]

## 2.2. Generatív modellek térnyerése

A generatív AI modellek, például a Generatív Adverzális Hálózatok (GAN) egyre nagyobb szerepet kapnak a digitális képfeldolgozásban. A GAN egy speciális típusú mesterséges neurális hálózat, amelyeket generatív modellezésre használnak. A GAN-ek célja, hogy új, realisztikus adatokat generáljanak a meglévő adatok alapján, például képeket, szövegeket vagy hangokat. A GAN-ek alapötletét Ian Goodfellow és munkatársai vezették be 2014-ben, és azóta széles körben alkalmazzák a mesterséges intelligencia és a gépi tanulás különböző területein. [3]

A GAN-ek két versengő neurális hálózatot használnak: egy generátort, amely képeket generál, és egy diszkriminátort, amely döntést hoz. A generátor felelős az új, szintetikus adatok generálásáért. A generátor véletlenszerű zajból (pl. egy véletlen vektorból) próbál új adatpontokat, például képeket előállítani, amelyek a lehető legjobban hasonlítanak az eredeti adatokra. A diszkriminátor felelős, hogy megkülönböztesse a valós adatokat a generátor által készített hamis adatoktól. A diszkriminátor megpróbálja eldönteni, hogy egy adott adatpont "valódi" (az eredeti adathalmazból származik) vagy "hamis" (a generátor által készített).

Ezt a technikát nemcsak képek szintetizálására használják, hanem arra is, hogy egy meglévő képen változtatásokat végezzenek (például zaj eltávolítása, felbontás növelése, stb.). A két hálózat közötti "játék" adja a GAN-ek működésének lényegét: a generátor folyamatosan javítja magát annak érdekében, hogy olyan adatokat készítsen, amelyeket a diszkriminátor nehezen tud megkülönböztetni a valós adatoktól, míg a diszkriminátor folyamatosan javul abban, hogy minél jobban felismerje a hamis adatokat.

A GAN-ek működését legjobban egy "versenyképes tanulási folyamatként" lehet leírni, ahol a generátor és a diszkriminátor egymással szemben próbál jobban teljesíteni.

A GAN-ek célja egy Nash-egyensúly elérése a generátor és a diszkriminátor közötti játékelméleti folyamatban. A GAN-ek egy veszteségi függvényt használnak, amely a két hálózat teljesítményét méri. A diszkriminátor célja a valós és a hamis adatok közötti különbség maximalizálása, míg a generátor célja, hogy minimalizálja ezt a különbséget, vagyis olyan adatokat generáljon, amelyeket a diszkriminátor nem tud helyesen osztályozni. GAN-okkal valósághű képeket lehet generálni, amelyeket például a játékokban, vizuális effektekben és tartalomgyártásban alkalmaznak, a képek egyik stílusból másikba alakíthatók, vagy éppen javíthatók, például felbontás növelése, elmosódás megszüntetése vagy zajcsökkentés révén.

## 2.3. Autoenkóderek

Az autoenkóderek egyfajta neurális hálózatok, amelyek célja a képek tömörítése, azaz a képekből olyan reprezentációkat készítenek, amelyek kevesebb információt tartalmaznak, mégis vissza tudják állítani az eredeti képet. Ezeket gyakran használják képjavításra vagy zajszűrésre, például egy kép zajos verziójából képesek tisztább verziót generálni. [7]

Az autoenkóder egy két részből álló hálózat. A kódoló (encoder) a bemeneti adatokat (például egy képet) egy kisebb dimenziójú térbe vetíti át, amelyet gyakran neveznek "rejtett térnek" vagy "latent space"-nek. Az encoder célja, hogy az eredeti adat fontos jellemzőit megtalálja és sűrítse. A dekódoló (decoder) a kódolt információt visszaalakítja az eredeti bemeneti adathoz hasonló formára. Ez a folyamat egy rekonstrukciós feladat, amelyben a hálózat megpróbálja visszanyerni az eredeti bemeneti adatot a rejtett reprezentáció alapján. A hálózat úgy van kiképezve, hogy minimalizálja a különbséget az eredeti bemeneti kép és a dekódolás után kapott kép között, tehát a rekonstrukciós hibát. A bemeneti kép (például egy 28x28 pixeles kézírásos számjegy) átmegegy a kódolón, amely egy tömörített reprezentációt hoz létre a képről. A kódoló egy kisebb dimenziójú vektort generál, amely a kép legfontosabb jellemzőit tartalmazza. Ezt nevezik a "rejtett térnek". A dekódoló

ezt a tömörített reprezentációt felhasználva megpróbálja visszaállítani a bemeneti képet. Az eredmény egy rekonstrukciós kép, amely hasonlít az eredeti bemenetre.

Az autoenkóder alkalmazásai:

- Dimenziócsökkentés: Az autoenkóderek képesek a képadatok dimenzióját csökkenteni azáltal, hogy megtalálják a legfontosabb jellemzőket. Ez hasznos lehet nagy adathalmazok kezelésénél, például képosztályozás vagy arcfelismerés során.
- Zajszűrés (denoising): Az úgynevezett denoising autoencoder változata arra van kiképezve, hogy zajos képekről megtanulja visszaállítani az eredeti, zajmentes képeket. Az ilyen autoenkódereket arra trenírozzák, hogy felismerjék az alapvető jellemzőket, és képesek legyenek eltávolítani a képen lévő zajt.
- Generatív modellezés: Az autoenkóderek egy speciális típusa, a variációs autoencoder (VAE), generatív modellek kialakítására használható. Ez azt jelenti, hogy nemcsak képesek egy képet rekonstruálni, hanem új, hasonló képeket is előállítani. Például egy VAE megtanulhat egy adott képstílust, és új képeket generálhat hasonló jellemzőkkel.
- Anomália detektálás: Az autoenkódereket gyakran használják anomáliák felismerésére is. Ha a hálózatot arra tanítják, hogy normál adatokat (például tipikus képeket) rekonstruáljon, akkor a szokatlan vagy hibás képeket nem tudja jól visszaállítani, és az ezekre vonatkozó rekonstrukciós hiba nagy lesz. Ez segíthet felismerni az adatbázisban található anomáliákat.

Az autoenkóder egy erőteljes eszköz a digitális képfeldolgozásban, előnye a felügyelet nélküli tanulás, azaz, hogy nem szükséges annotált adat, mivel az autoenkóder maga az adatból tanulja meg a fontos jellemzőket. A kódoló réteg képes jelentős mértékben csökkenteni az adat dimenzióját, miközben megőrzi annak lényeges tulajdonságait. Ez a modell alapvető része sok modern képfeldolgozó alkalmazásnak, mint például a képi adatbázisok optimalizálása, anomália detektálás, valamint képminőség javítás.

## 2.4. Szupport vektor gépek

A szupport vektor gép (Support Vector Machine, SVM) egy felügyelt tanulási algoritmus, amelyet elsősorban osztályozási és regressziós feladatok megoldására használnak. Az SVM az egyik legismertebb és legszélesebb körben használt gépi tanulási algoritmus, különösen akkor, ha az adatok jól elkülöníthetőek.

Az SVM-ek a gépi tanulási algoritmusok egy korábbi, klasszikus megközelítését képviselik, amelyeket széles körben használtak képosztályozási feladatoknál. Bár manapság a CNN-ek és mélytanulási módszerek váltak dominánssá, az SVM-ek még mindig használhatók olyan feladatokra, ahol kevés adat áll rendelkezésre, és a feladat nem túl komplex. Az SVM alapötlete, hogy a bemeneti adatokat (pl. képeket, szövegeket stb.) egy magasabb dimenziójú térbe vetíti, és megpróbál egy hipersíkot találni, amely a legjobban elkülöníti a különböző osztályokat. A hipersík a tér egy olyan része, amely a két osztály közötti lehető legnagyobb távolságot (margin-t) hozza létre. A szupport vektorok azok az adatpontok, amelyek a legközelebb helyezkednek el a hipersíkhhoz, és meghatározzák annak helyzetét. A többi adatpont nem befolyásolja a hipersík helyzetét, csak ezek a szupport vektorok.

Amikor az adatok lineárisan elválaszthatóak, azaz van olyan egyenes (vagy magasabb dimenzióban hipersík), amely pontosan el tudja választani a két osztályt, a lineáris SVM jól működik. Sok valós problémában az adatok nem lineárisan elválaszthatóak. Ilyen esetekben az SVM egy kernel trükköt használ, amely az adatokat egy magasabb dimenziójú térbe vetíti át, ahol lineárisan elválaszthatóvá válnak.

Az SVM gyakorlatilag egy optimalizálási problémát old meg: minimalizálja a veszteségi függvényt úgy, hogy maximalizálja a margin-t a két osztály között. Ha az adatok nem teljesen lineárisan elválaszthatóak, akkor az SVM egy bizonyos fokú hibát megenged, amit a soft margin fogalma ír le. Ebben az esetben egy szabályozási paraméter segít kiegyensúlyozni a hibás osztályozás és a margin maximalizálása közötti kompromisszumot.

## 2.5. Hibrid modellek alkalmazása

A képfeldolgozásban egyre inkább teret nyernek a hibrid megoldások, amelyek különböző AI technikák, amelyek a hagyományos képfeldolgozás és a szabályalapú rendszerek kombinációjára épülnek. Ezáltal robusztusabb és megbízhatóbb rendszerek alakíthatók ki.

## 3. AZ AI ALAPÚ KÉPFELDOLGOZÁS ALKALMAZÁSI TERÜLETEI

### 3.1. Reálisabb és élethűbb képgenerálás

A generatív modellek fejlődésével egyre reálisabb és részletgazdagabb képek hozhatók létre AI segítségével.

- **Fotórealisztikus képek előállítása:** A legújabb AI modellek képesek olyan fotórealisztikus képeket generálni, amelyeket nehéz megkülönböztetni a valós fotóktól. Ezeket a képeket például a reklámpia, a filmgyártás vagy az online tartalomszolgáltatók használhatják.
- **Mesterséges arcok és testek generálása:** Az AI már képes valóság-hű emberi arcokat és karaktereket generálni, amelyeket videójátékokban, filmekben vagy akár a virtuális valóságban is alkalmaznak.

### 3.2. Orvosi képfeldolgozás forradalmasítása

Az AI-alapú képfeldolgozás az orvosi diagnosztikában is gyors fejlődést mutat.

- **Automatizált diagnosztika:** A deep learning algoritmusok segítségével automatikusan felismerhetők különböző betegségek, mint például a daganatok, érrendszeri problémák vagy retinabetegségek. A gépi tanulás alapú módszerek gyorsabbak és pontosabbak lehetnek, mint az emberi elemzés.
- **Képszegmentálás és elemzés:** Az orvosi képeken történő szegmentálás, például MRI vagy CT felvételek esetében, lehetővé teszi az egyes szövetek, szervek vagy elváltozások pontosabb azonosítását.

### 3.3. Valós idejű képfeldolgozás

Az AI-alapú rendszerek egyre inkább valós idejű képfeldolgozást tesznek lehetővé, ami kulcsfontosságú a biztonsági rendszerekben, az autonóm járművekben és a drónok irányításában.

- **Autonóm járművek:** Az AI képes valós időben feldolgozni a környezetet, és segíti a járművet a navigálásban, felismeri az akadályokat és biztosítja az ütközések elkerülését.
- **Arcfelismerés:** Valós idejű arcfelismerés terjedése biztonsági rendszerekben és közösségi médiában, például személyazonosításhoz vagy beléptető rendszerekhez.

### 3.4. Edge AI és beágyazott rendszerek

Az Edge AI lehetővé teszi a képfeldolgozás decentralizált alkalmazását, például okoskamerákban, IoT-eszközökben és beágyazott rendszerekben. Az edge computing és AI kombinációja lehetővé teszi, hogy a képfeldolgozás az eszközökön magukon történjen, ezáltal csökkentve a késleltetést, és gyorsabb reakciót biztosítson például biztonsági kamerarendszereknél vagy gyártósoroknál.

### 3.5. Interaktív és személyre szabott képfeldolgozás

Az AI képes a felhasználók viselkedésének és preferenciáinak tanulmányozására, hogy egyedi vizuális tartalmat hozzon létre. A képfeldolgozó rendszerek alkalmazhatók arra, hogy a felhasználók vizuális igényeihez igazodó, testreszabott képeket vagy grafikai elemeket hozzanak létre például e-kereskedelmi platformokon vagy hirdetési rendszerekben.

## 4. ÖSSZEFOGLALÁS

Az AI alapú digitális képfeldolgozás átalakította a képelemzés és -manipulálás módszereit, lehetővé téve a valós idejű objektumfelismerést, képszegmentálást és még a hiányos képek helyreállítását is. A mesterséges intelligencia lehetővé teszi a képek és videók valós idejű feldolgozását, ami azonnali elemzést és döntéshozatalt tesz lehetővé (pl. felügyeleti rendszereknél valós idejű megfigyelés és a gyanús tevékenységek észlelése, vagy élő közvetítésnél a videóminőség javítása és szűrők alkalmazása valós időben). A mélytanulási modellek, különösen a CNN-ek és GAN-ek, forradalmasították az orvosi képalkotástól a kreatív iparágakig terjedő alkalmazásokat. Az AI-alapú digitális képfeldolgozás területén a legújabb trendek közé tartoznak a

deep learning modellek alkalmazása, a generatív modellek fejlődése, az orvosi képfeldolgozás fejlődése és a valós idejű alkalmazások növekedése.

A legmodernebb AI alapú képfeldolgozási technikák szinte teljes mértékben mély neurális hálózatokra építenek, különösen a konvolúciós neurális hálózatokra, amelyek kifejezetten a képadatok feldolgozására lettek kifejlesztve. A CNN-ek képesek automatikusan kivonni a fontos jellemzőket a képekből, amelyeket a hálózat megtanul, így nincs szükség kézi jellemzők kiválasztására.

A Generatív Adverszárius Hálózatok képesek új, valóságghű képeket létrehozni meglévő adatok alapján, vagy olyan módon manipulálni képeket, ahogyan azt egy hagyományos algoritmus nem tudná. A gépi tanuláshoz és a neurális hálózatoknak köszönhetően az AI-módszerek automatizálják és pontosabban végzik el az összetett elemzéseket, mint a hagyományos képfeldolgozás. Széles körben használják az egészségügyben a diagnosztikában, a biztonsági rendszerekben a megfigyeléshez és a szórakoztatásban a digitális élmények fokozásához.

A modern arcfelismerő rendszerek mélytanulási modelleket használnak, hogy felismerjék az egyedi arcokat és azok jellemzőit, akár a tömegből is. Az autonóm járművek és robotok is nagyban építenek a gépi tanulásra a képfeldolgozásban. A mesterséges intelligencia alapú képfeldolgozás segíti az autonóm járművek navigációját és a környezet érzékelését. A járművek képesek felismerni az útvonalakat, közlekedési táblákat, akadályokat, vagy akár a gyalogosokat is, amelyek alapján döntéseket hoznak.

Ugyanakkor az AI technológiák használata új kihívásokat is felvet, például az etikai kérdéseket és a hamisított tartalmak elleni védekezést illetően. A deepfake technológiával készült képek és videók visszaéléseket okozhatnak, ezért az AI másrészt éppen ezen hamisítványok felismerésére és azonosítására törekszik.

Összességében az AI képfeldolgozás és a dokumentum-központú munkafolyamatok könnyen optimalizálhatók automatizálással az adatelemzés pontosságának javítása és a szervezeteken belüli zökkenőmentes együttműködés elősegítése érdekében. A gépi tanulási rendszerek nagy mennyiségű képadatot képesek feldolgozni és elemezni rövid időn belül, ami lehetővé teszi az alkalmazásukat ipari szinten is.

## IRODALMI HIVATKOZÁSOK

- [1] Wilhelm Burger, Mark J. Burge *Digital Image Processing*. Springer Cham, 2022, eBook ISBN 978-3-031-05744-1
- [2] Solanki, S. R., Khublani, D. K. *Generative Artificial Intelligence*. Springer, Berlin, 2024, EAN 9798868804021
- [3] Foster, D. *Generative Deep Learning*. O'Reilly Media, 2023, ISBN 1098134184
- [4] Simran, A., Shijin Kumar, P. S., & Bachu, S. *Content based image retrieval using deep learning convolutional neural network*. IOP Conference Series: Materials Science and Engineering, 2021, 1084(1), 012026, <https://doi.org/10.1088/1757-899X/1084/1/012026>
- [5] Li, Y., & Wang, M. *Image retrieval algorithm based on deep learning*. In ICGSP 2020: 2020 The 4th International Conference on Graphics and Signal Processing. Association for Computing Machinery.
- [6] Kherraki, A., & El Ouazzani, R. *Deep convolutional neural networks architecture for an efficient emergency vehicle classification in real-time traffic monitoring*. IAES International Journal of Artificial Intelligence, 2022, 11(1), 110–120.
- [7] Cheng-Yu Chen, Jenq-Shiou Leu and Setya Widyawan Prakosa *Using Autoencoder to Facilitate Information Retention for Data Dimension Reduction*, IEEE, pp. 1-5, 2018.