

Kiss Csaba¹

A katonai gyakorlóruhán lévő rendfokozati jelzések valós idejű felismerése gépi látás és mesterséges intelligencia segítségével

Ma már megszokott, hogy a civil világban a korszerű okostelefonok és tabletek elterjedése miatt egyre inkább megjelennek olyan alkalmazások, amelyek a valós életben látottakat többletinformációval egészítik ki. Ennek alapján ezek az eszközök a kamerakép segítségével felismerik a megadott objektumot, majd annak megfelelően a valóságot virtuálisan kibővítik. A dolgozat célja bemutatni a katonai gyakorlóruhán viselhető tépőzárás rendfokozati jelzések felismerési folyamatát. Ennek során az egyes módszerek elméleti háttérének ismertetésén túl kitérek az OpenCV gépilátás-könyvtár alkalmazására, valamint az erőforrásigény vizsgálatára. A publikációban három főbb objektum felismerésre alkalmas módszert vizsgállok meg: a korrelációs mintaillesztés-alapú felismerést, a kaszkádosított osztályozó alapú felismerést, illetve a jellegzetességmegfeleltetés alapján való felismerést.

Kulcsszavak: mesterséges intelligencia, gépi látás, rendfokozati jelzések

REAL-TIME RECOGNITION OF RANK INSIGNIA ON MILITARY UNIFORMS USING MACHINE VISION AND ARTIFICIAL INTELLIGENCE

It is now commonplace that, due to the spread of modern smartphones and tablets in the civilian world, applications that supplement what you see in real life with additional information are increasingly appearing. Based on this, these devices use the camera image to recognize a specified object, and accordingly, reality is virtually expanded. The aim of the publication is to present the recognition process of the Velcro rank markings worn on military training clothes. In doing so, in addition to explaining the theoretical background of each method, I will cover the application of the OpenCV machine vision library and the examination of resource requirements. In the publication, I examine three main methods

¹ BriCS Szoftverfejlesztő és Tanácsadó Kft., infokommunikációs vezető, e-mail: kiss.csaba@uni-nke.hu; ORCID ID: 0000-0002-7265-8704.



suitable for object recognition: recognition based on correlation pattern matching, recognition based on cascaded classifiers, and recognition based on feature matching.

Keywords: *artificial intelligence, machine vision, rank signals*

Bevezetés

A nemzetközi trendekkel párhuzamosan Magyarországon is elindult a mesterséges intelligencia (a továbbiakban: MI) egy fejlesztési folyamata, amelynek során az Innovációs és Technológiai Minisztérium (a továbbiakban: ITM) vezetésével létrehozták a Mesterséges Intelligencia Koalíciót (a továbbiakban: MIK), amelynek elsődleges célja a hazai MI-technológiák nemzetközi versenyképességének javítása, beépítésük az üzleti életbe és az állami folyamatokba.²

Az MI felhasználása a haderőben minden területen elindult, ahogy a civil életben is. A publikáció az MI felhasználásának egyik lehetőségét mutatja be a katonai rendfokozatok felismertetésével.

A mai értelemben vett katonai rendfokozatok kialakulása és a katonai hierarchia megjelenése a 17–18. században a tömeghadseregek létrejöttének időszakára tehető, azonban a rendfokozati jelzések rendszeresítése csak hosszú évek alatt valósult meg. A magyar rendfokozati megnevezésekben elsősorban a német, illetve a francia hatás érződik.

A jelenleg is hivatalos rendfokozatokat a honvédek jogállásáról szóló 2012. évi CCV. törvény I. melléklete sorolja fel.³ Ezek a rendfokozati jelzések 1999. március 12-én, Magyarország NATO-taggá válásának napján léptek érvénybe, főbb jellemzőik: zöld alapszín, minden rendfokozatban himzett szegély és himzett hatágú csillag.

A katonai kiképzés során a katonák a magyar rendfokozati jelzéseket is megtanulják. Mivel Magyarország a NATO tagja, és ebből adódóan a közös katonai gyakorlatokon kapcsolatban vagyunk külföldi katonákkal, így az ő rendfokozati jelzésüket is illő ismerni. A NATO-nak már 31 tagországa van. Ez azt jelenti, hogy 31 ország hadseregének a különböző színű és alakú rendfokozati jelzéseit illene felismerni. A katonák közötti együttműködés alapja a katonai tiszteletadás, amelyet többek között a rendfokozati jelzés szerint adunk meg, ezért is fontos ennek az ismerete. Előfordulhat olyan külföldi országok haderejében rendszeresített

² PORKOLÁB–NÉGYESI 2019.

³ 2012. évi CCV. törvény a honvédek jogállásáról.

rendfokozat, amely a magyar rendszerből hiányzik (haditengerészet, űrhaderő, tengerészgyalogság), viszont a külföldi missziók és gyakorlatok alatt fontos ezek ismerete is. Ennek az ismeretnek a megszerzésében segítheti a katonát egy gépi látáson alapuló rendfokozati jelzéseket értelmező folyamat.

A publikációban ismertetett folyamat eredményeképpen más katonai objektumok, eszközök felismertetése is lehetővé válhat. A gépi látást a haderőben több helyen is lehet alkalmazni, elsősorban olyan területeken, ahol a tárgyak, eszközök vagy jelzések sokszínűsége egyszerre fordul elő. A gépi látás segítségével lehetőség nyílik a fegyverek felismertetésére, a lőszer megkülönböztetésére, a technikai eszközökön lévő jelzések értelmezésére stb. A gépi látás útján megszerzett gyors ismeret a katonai eszközökről vagy katonai jelzésekről növelheti a katona magabiztosságát, ezenkívül szélesítheti vagy megerősítheti tudását. A katonai műveletekben részt vevő katonák életét mentheti meg egy felülről támadó drón korai észrevétele, amit gépi látás útján lehet elősegíteni.

A tanulmány célja egy gépi látáson alapuló, a katonai rangjelzések felismerését támogató folyamat ismertetése. A rendfokozati jelzések által nyújtott szín- és formavilág felismerése és értelmezése tipikusan gépi látási feladat. A gépi látás a gép által a környezetéről, az abban lévő objektumokról készített kép számítógépes elemzését jelenti abból a célból, hogy a készített képből az ember számára meghatározott információt nyerjen ki, vagy pedig a saját működéséhez szükséges információt megszerezze.

A gépi látással kialakított rendszerek a modern technológiának köszönhetően képesek lehetnek különféle feladatok ellátására, mint például az arcfelismerés, az autonóm járművek navigálása, az iparban minőségellenőrzési feladatok végrehajtása, az orvoslásban a képelemzési funkciók segítése. A gépi látás használata hatékonyabbá teheti a folyamatokat és csökkentheti az emberi beavatkozás szükségességét.

A gépi látás kialakítása a *deep learning* technikák és az általános neurális hálózatok használatával érhető el. Az ilyen típusú gépi tanuló algoritmusok az emberi agyhoz hasonlóan működnek, ami lehetővé teszi a számítógépeknek, hogy automatikusan tanuljanak, és finomítsák látási képességeiket az idő múlásával.

A gépi látást használó rendszerek alkalmazási lehetősége nagyon változatos, és különösen akkor lesz széles, ha általánosan elterjednek, mert lehetőség nyílik rá, hogy egyszerűbb eszközök is használhassák ezt a technológiát.⁴

⁴ KOLLÁR–NAGY 2021.

Gépi látás

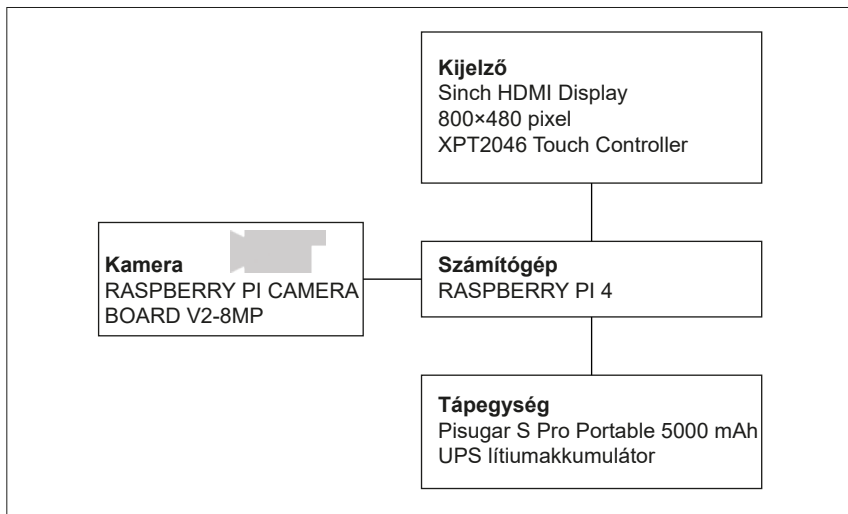
A gépi látás a számítógépes látástechnika egyik ága, amely lehetővé teszi, hogy számítógépes rendszerek segítségével feldolgozzák és értelmezzék a digitális képeket. A gépi látás az MI képfeldolgozási algoritmusait alkalmazza az objektumok azonosítására, amelyeket a számítógép- vagy kameraképeken észlel. A gépi látás alatt a teljes technikai folyamatot értjük, a felhasználási cél eléréséig. Ennek megfelelően a rendszer elemei a következők lehetnek:

1. *képalkotás* – ez jellemzően egy kamera;
2. *jelfeldolgozás* – vizuális adat feldolgozására használatos célhardver, például videokártya;
3. *szoftver* – adatok értelmezéséhez szükséges speciális programok;
4. *kommunikáció* – az eredmények megjelenítése.

Képalkotó berendezésnek nevezzük a mért fizikai jellemzőhöz illeszkedő olyan hardvereszközt, amely a mért adatokból képet vagy képek sorozatát képes előállítani. A képalkotó berendezések a gépi látásban nagyon sokfélék lehetnek, mivel nagyon sok fizikai jellemzőből van lehetőség képet előállítani. A képalkotó berendezések csoportosítása több szempont szerint is lehetséges, beszélhetünk dimenzionális felosztásról, felépítésalapú felosztásról, elhelyezési felosztásról, élettani felosztásról.

A *jelfeldolgozó egység* (hardver és szoftver) együttese végzi a digitális képfeldolgozási és képelemzési folyamatot. A képfeldolgozás alatt ebben a folyamatban olyan komponenseket, részegységeket értünk, amelyek képből képi információt állítanak elő. A képfeldolgozás körébe tartozik egyebek között a videók kódolása, a színtér-transzformációk végrehajtása, a képkivágás, a színszűrés, az éldetektálás stb. A képelemzés alatt pedig olyan komponenseket értünk, amelyek képből nem képi jellegű adatok előállítását célozzák. Idetartozik például a képi folt területének meghatározása, a mozgáselemzés, a mozgásirány becslés, a pozíció meghatározása, a mélységbecslés, a jellemző pontok kinyerése, a sajátértékek és vektorok kinyerése stb.

A *szoftver* a gépi látás folyamatának kulcseleme, az összegyűjtött jeleket értelmező, kontextusba helyező döntési logika, amely egy programkód. Ez lehet mély tanulást végző neurális háló vagy programozható elektronikus áramkör vagy programozható logikai vezérlőre (*Programmable Logic Controller; PLC*) írt döntési logika. Általában a képalkotó berendezést gyártó cég a felhasználók részére biztosít előre megírt, általános célú vagy az eszköz által megcélzott



1. ábra. Hardveres kiépítési lehetőség

Forrás: a szerző szerkesztése

specifikus tevékenység ellátásához szükséges programozható függvénykönyvtárat és segédprogramokat. Az elérhető, fontosabb szoftverkönyvtárak:

- a) AForge.NET;
- b) OpenCV (Open Computer Vision Library);
- c) VLFeat;
- d) PCL (PointClouds).

A *kommunikációhoz* a döntés–értelmezés–felismerés folyamatban létrejövő eredmények közlésére vagy vezérlésre alkalmas hardver- és szoftverkomponenseket soroljuk.

A gépi látás rendszerelemeire megvalósuló hardveres kiépítési lehetőséget mutatja az 1. ábra.

Az 1. ábra szerint a képalkotást a *Kamera*, a jelfeldolgozást és -elemzést a *Számítógép* végzi, kommunikációként pedig a *Kijelző* funkcionál.

Mivel ismerjük a gyakorlóruhán elhelyezett rendfokozati jelzés előlnézeti helyét az archoz képest, ezért a katonai rendfokozatok felismertetését azzal lehet gyorsítani, hogy egy arcfelismerő segítségével megjelöljük az arc helyét a képen, így az archoz viszonyítva könnyen behatárolhatjuk a rendfokozat helyét, azaz



2. ábra. Gyakorlórúhán hordható rendfokozati jelzések

Forrás: a szerző szerkesztése

kivághatjuk a forrásképből. Ennek az az előnye, hogy kisebb képet kell majd analizáltatnunk a szoftverekkel, ami a felismerésben nagyságrendi időbeni csökkenést eredményez. A másik előnye, hogy kisebb teljesítményű számítógépeken is rövid idő alatt eredményt kaphatunk. Ilyen számítógép lehet a Raspberry Pi, a HP T520, a NanoPi, az Orange Pi, a Banana Pi, a Geniatech XPI.

A katonai rangjelzések felismertetése speciális eset, mert a rangjelzések színeit már ismerjük, így azokra előszűrést lehet végezni. Mivel a katonai rangjelzések színe specifikusan csak a rangjelzésekre jellemző, sem az állatok között, sem a növényi világban nem fordul elő, így ezeknek a színeknek a detektálása egy képterületen azt jelentheti, hogy ott rendfokozati jelzés található. A következő fejezet a színek detektálását mutatja be.

A katonai rendfokozatok értelmezése gépi látáson keresztül

A katonai rendfokozat a katona szolgálati rendben elfoglalt helyét meghatározó besorolás. A rendfokozat megjelenítésére szolgál a rendfokozati jelzés, amely lehet váll-lap, tépőzárás (gyakorlórúhán), paroli, sapkajelzés, karrendfokozati jelzés.

Mivel a katona a harctéren gyakorlórúhában teljesít szolgálatot, ezért jelen tanulmányban a gyakorlórúhán hordható magyar tépőzárás rendfokozati jelzés felismeréséről lesz szó, nappali fényviszonyok között.

A Magyar Honvédségben tisztések, altiszték, zászlósok, tiszték, főtiszték, tábornokok szolgálnak, ezért a rendfokozati jelzések ezen rendfokozati különbségek



3. ábra. Rendfokozati jelzések felépítése
a színek felosztásában

Forrás: a szerző szerkesztése

egyértelmű megkülönböztetésére szolgálnak. Ezt a megkülönböztetést színben, formában és a csillagok számában mutatja meg a rendfokozati jelzés (2. ábra).

A rendfokozati jelzések felépítésénél észrevehetünk markánsan elkülönülő színeket, amit a 3. ábra mutat.

A rendfokozati jelzések feldolgozhatósága érdekében rendezzük őket csillagszám szerint (1. táblázat).

1. táblázat. Rendfokozati jelzések csillagszám szerint

A csillagok száma	Honvéd	Tisztesek	Alttisztek	Zászlósok	Tisztek	Főtisztek	Tábornokok
4							vezér- ezredes
3		szakasz- vezető	főtörzs- őrmester	főtörzs- zászlós	százados	ezredes	altábornagy
2		tizedes	törzs- őrmester	törzszászlós	főhadnagy	alezredes	vezér- őrnagy
1		örvezető	őrmester	zászlós	hadnagy	őrnagy	dandár- tábornok
Nincs	honvéd						

Forrás: a szerző szerkesztése

Nézzük meg a színek előfordulását a rendfokozati jelzéseken az 1. táblázat szerinti felosztásban (2. táblázat)!

2. táblázat: *Rendfokozati jelzések szín szerint*

<i>Szín</i>	<i>Honvéd</i>	<i>Tisztesek</i>	<i>Altisztek</i>	<i>Zászlósok</i>	<i>Tisztek</i>	<i>Főtisztek</i>	<i>Tábornokok</i>
<i>Alapszín</i>	barna	barna	barna	barna	barna	barna	nincs
<i>Csillag-alapszín</i>	nincs	nincs	nincs	nincs	nincs	világos-barna	világos-barna
<i>Csikszín</i>	nincs	nincs	világos-barna	világos-barna	nincs	nincs	nincs
<i>Csillagszín</i>	nincs	fehér	világos-szürke	sárga	sárga	világos-szürke	világos-szürke

Forrás: a szerző szerkesztése

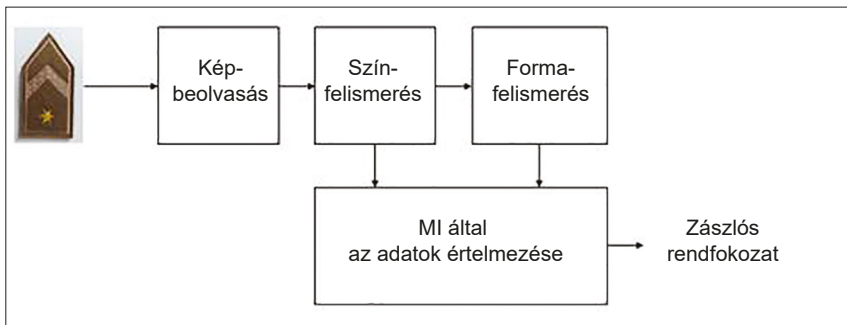
A 2. táblázat szerint látható, hogy a rendfokozati jelzések öt különböző szín segítségével épülnek fel. Ezek a színek a következők: fehér, sárga, világosbarna, barna és világosszürke. A gépi látás a rendfokozati jelzéseken lévő színeket detektálni tudja, azaz 1 és 0 értéket ad a szín változójához, így ha talál olyan színt, az 1 értéket, ha nincs olyan szín, akkor a 0 értéket (3. táblázat).

3. táblázat. *Színek kódolása a rendfokozati jelzéseken*

<i>Szín</i>	<i>Honvéd</i>	<i>Tisztesek</i>	<i>Altisztek</i>	<i>Zászlósok</i>	<i>Tisztek</i>	<i>Főtisztek</i>	<i>Tábornokok</i>
<i>Fehér</i>	0	1	0	0	0	0	0
<i>Sárga</i>	0	0	0	1	1	0	0
<i>Világos-barna</i>	0	0	1	1	0	1	1
<i>Barna</i>	1	1	1	1	1	1	0
<i>Világos-szürke</i>	0	0	1	0	0	1	1

Forrás: a szerző szerkesztése

A 3. táblázatból jól kivehető, hogy az altisztek és a főtisztek színekódolása megegyezik, a többi különbözik egymástól. Kimondhatjuk, hogy csak színekódolásból



4. ábra. A képfeldolgozás folyamata

Forrás: a szerző szerkesztése

nem lehet az altisztet a főtiszttól megkülönböztetni. A különbséget a 2. táblázat tartalmazza, ahol látható, hogy az altisztek rendfokozati jelzése tartalmaz egy csíkot, ebből következik, hogy a színek mellett a rendfokozati jelzések formáját is meg kell határozni a helyes felismeréshez (4. ábra).

A szín megállapításánál fontos tényező, hogy a különböző fényviszonyoknál készült kép mennyire lesz feldolgozható. Az erősen fényvisszaverő felület csillogást okozhat, ami sokszor felismerhetetlenné teszi a képrészletet vagy a képet, így a szín felismerése is lehetetlenné válik. A gyakorló egyenruha álcázást is szolgál, így a rajta lévő katonai rendfokozati jelzések is nem „erősen” fényvisszaverő anyagból készülnek, ezért a képek is jól feldolgozhatók.

A színek felismerése RGB (Red–Green–Blue, vörös–zöld–kék) -színkódok alapján történik, Michael Shneier módszere szerint. A gyakorlatban ez azt jelenti, hogy a kép egy pixelének színét függvény segítségével számmá alakítjuk. A függvény az RGB-kód minimum- és maximumértékeit figyeli az 5. ábra szerint.

$$m(x,y) = \begin{cases} k_1 & \text{ha} \begin{cases} R_{\min} \leq f_r(x,y) \leq R_{\max} \\ G_{\min} \leq f_g(x,y) \leq G_{\max} \\ B_{\min} \leq f_b(x,y) \leq B_{\max} \end{cases} \\ k_2 & \text{másik szín} \end{cases}$$

5 ábra. RGB-küszöbölés

Forrás: SHOJANIA 2003

Az $m(x,y)$ függvény adja vissza az (x,y) koordinátájú képpont egyes összetevőit $f_{r,g,b}(x,y)$ a forrásképből egy számba kódolva (például k_i). Mivel a $R_{\min}$, a $G_{\min}$, a $B_{\min}$, az $R_{\max}$, a $G_{\max}$ és a $B_{\max}$ értékét a katonai rangjelzésen található színek alapján állítjuk be, így a k_1 érték jelentheti például a sárga szín meglétét, a k_2 akár a világosbarnáét. A 3. táblázatban megszámlolva a színek előfordulását a 4. táblázatot kapjuk.

4. táblázat. *A színek előfordulása*

<i>Szín</i>	<i>Előfordulási mennyiség</i>
Fehér	1
Sárga	2
Világosszürke	3
Világosbarna	4
Barna	5

Forrás: a szerző szerkesztése

Az előfordulási mennyiség határozza meg a súlyozást, amely egymástól számokban elkülöníti a rendfokozati jelzéseket. A legtöbbször előforduló szín kapja a legkisebb súlyozási értéket és a legkevésbé előforduló szín kapja a legnagyobb súlyozási értéket. A 4. táblázatot kiegészítve súlyozási értékekkel (kettes számrendszert használva) kapjuk meg az 5. táblázatot.

5. táblázat. *Súlyozási értékek*

<i>Szín</i>	<i>Előfordulási mennyiség</i>	<i>Súlyozási érték</i>
Fehér	1	$(2^4)=16$
Sárga	2	$(2^3)=8$
Világosszürke	3	$(2^2)=4$
Világosbarna	4	$(2^1)=2$
Barna	5	$(2^0)=1$

Forrás: a szerző szerkesztése

A kettes számrendszer előnye, hogy a színek behatárolása és átkódolása, majd összegzése után minden rendfokozat csak egy, rá jellemző számértéket kap. Amennyiben figyelembe vesszük a 30 különböző ország rendfokozati jelzéseit és a felhasznált színeket, akkor összegzés után minden ország a rendfokozati

jelzéseire jellemző – feltehetően – egyedi számot fog kapni. Amennyiben mégis egyezne a kapott szám, mert az országok ugyanazt a szint használják a jelzéseikre, akkor a formának a beazonosításával tovább lehet finomítani a felismerést. A súlyozási érték szerint átírjuk a 3. táblázatot, és összeadjuk a kapott értékeket (6. táblázat).

6. táblázat. Összegzett súlyozási értékek

<i>Szín</i>	<i>Honvéd</i>	<i>Tisztesek</i>	<i>Altisztek</i>	<i>Zászlósok</i>	<i>Tisztek</i>	<i>Főtisztek</i>	<i>Tábornokok</i>
Fehér	0	16	0	0	0	0	0
Sárga	0	0	0	8	8	0	0
Világosbarna	0	0	2	2	0	2	2
Barna	1	1	1	1	1	1	0
Világosszürke	0	0	4	0	0	4	4
Összesen	1	17	7	11	9	7	6

Forrás: a szerző szerkesztése

A 6. táblázatból jól kivehető, hogy a rendfokozati jelzések számokban is elkülönülnek egymástól, kivéve az altisztek és a főtisztek, mert ők ugyanazt a számértéket kapták. Tehát a magyar rendfokozati jelzések ebben az esetben az 1–17 számértéket kapták. Ha például nézzük az angol szárazföldi rangjelzéseket, akkor azt látjuk, hogy megjelenik a piros szín. A piros fogja kapni a 2⁵ értéket. Tehát azoknak a rendfokozati jelzéseknek az esetében, ahol ez a piros szín is előfordul, a kapott számérték több lesz, mint 32. Az amerikai gyakorlóruhán megjelenik a sötétszürke, amely kaphatja a 2⁶ értéket. Tehát azoknak a jelzéseknek az esetében, ahol ez a szürke szín megtalálható, a kapott számérték több lesz, mint 64, és így tovább. Mivel minden ország törekszik a történelmi hagyományainak megfelelő színeket és formát választani a rendfokozati jelzéseinek, ezért nagy valószínűséggel elmondható, hogy a rendfokozatok szín szerinti kódolása országonként különböző értéket hoz. Természetesen vannak kivételek. Jelen esetben a magyar altisztek és főtisztek rendfokozati jelzéseinek számértéke megegyezik, ezért megkülönböztetésük a forma (objektum) felismerésnél lehetséges.

Objektumfelismerés

A képek alapján elvégzett automatikus objektumfelismerés a mai napig meglehetősen komplex feladat, amely nagy kihívás elé állítja a különböző képfeldolgozó rendszereket. A katonai rangjelzések felismerése általában könnyű az emberek számára, de ez a feladat a számítógépek számára igen nehéznek bizonyul. Fel akarjuk ismerni egy katona rangjelzését függetlenül a megvilágítástól, a kamerához viszonyított helyzetétől és a különböző hadseregeknél rendszeresített variációitól.

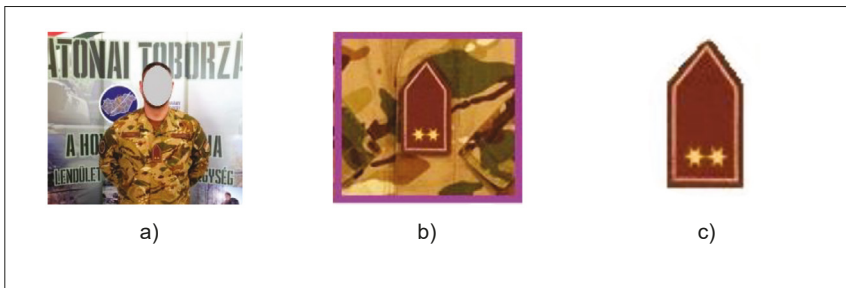
Ezen variációk bármelyike nagyarányú változásokhoz vezet a képpontok fényességértékében, ezért a képpontok közvetlen összehasonlítása a rangjelzések felismerése érdekében valószínűleg nem járható út. A mintaillesztés (*Template Matching*) az egyik legegyszerűbb módszer, amelyet objektumok felismerésére fel lehet használni. Egyszerű működése miatt azonban csak bizonyos, ideálishoz közeli körülmények mellett használható megfelelő hatékonysággal, igaz, összetettebb objektumok felismerésére nem kifejezetten alkalmas.

A mintaillesztés-alapú objektumfelismerés

A módszer működésének lényege az, hogy egy előre meghatározott mintaképhez viszonyítva a vizsgált kép részei mennyire illeszkednek, azaz mennyire hasonlók. A hasonlóság több korrelációt felhasználó módszer segítségével hatékonyan kiszámítható.

A mintaillesztés-alapú objektumfelismerés folyamata

Két képre van tehát szükség az illesztés elvégzéséhez: egy mintaképre (*T – Template Image*), amely tartalmazza a katonai rendfokozati jelzést, és egy forrásképre (*I – Source Image*), amelyen a keresést kell végezni. A módszer a mintakép hasonlóságának mértékét a forráskép egyes területeihez úgy határozza meg, hogy a forráskép egyik sarkából kiindulva a mintaképet pixelről pixelre vízszintesen és függőlegesen elmozgatja egészen addig, amíg a forrásképet teljesen be nem járta, és minden vizsgált helyen meg nem határozta a mintaképnek és a forráskép aktuálisan vizsgált részének a hasonlóságát.



6. ábra. a) Forráskép; b) a forrásképből kinyert katonai rendfokozati jelzés; c) mintakép
Forrás: a szerző szerkesztése

A módszerből következik, hogy a forrásképnek olyan paraméterekkel kell rendelkeznie szélességben és magasságban is, mint a mintaképnek. Ezt úgy érjük el katonai rendfokozati jelzések felismerésénél, hogy nem az egész forrásképen végzünk keresést, hanem csak a forrásképből kivágott, a keresendő jelzést tartalmazó képen (lásd 6/a–c. ábra).

A mintaillesztés-alapú objektumfelismerési módszer a forráskép egyes területeihez tartozó illeszkedések mértékét eredménymátrixban (R) tárolja el. Az eredménymátrixban tárolt értékek arra utalnak, hogy a forrásképből kinyert katonai rendfokozati jelzések mennyire hasonló a mintaképhez. A forrásképből kinyert katonai rendfokozati jelzések mérete: W – szélesség, H – magasság. A mintakép mérete: w – szélesség, h – magasság. Amennyiben a forráskép és a mintakép mérete sorban $W \times H$, illetve $w \times h$, az eredménymátrix mérete szintén a fenti egyszerű keresési módszerből adódóan $(W - w + 1) \times (H - h + 1)$. Az algoritmus lefutása után a legjobb illeszkedést az eredménymátrix globális minimuma vagy maximuma adja az illeszkedés meghatározásához felhasznált konkrét módszertől függően.

Ennél a módszernél figyelembe kell venni, hogy az összes katonai rendfokozati mintaképével el kell végezni a keresési folyamatot. Az 1. ábra alapján csak a magyar szárazföldi haderőknél lévő gyakorlóöltözetben viselhető rendfokozatok variációinak száma eléri a 20-at. Azonban, ha hozzávesszük a légierőt és a többi NATO-ország (31 ország) katonai rendfokozati jelzéseit is, akár többeszes nagyságú variációt kaphatunk, ami lassíthatja a felismerést. Amik viszont segíthetik a mintaillesztés-alapú objektumfelismerést, az *OpenCV*-ben rejlő tulajdonságok.

Az OpenCV által támogatott mintaillesztés-alapú objektumfelismerés

Az *Open Source Computer Vision Library (OpenCV)*⁵ olyan nyílt forráskódú *Berkeley Software Distribution (BSD)* licenccel rendelkező könyvtár, amely több száz számítógépes látási algoritmust tartalmaz. Az OpenCV moduláris felépítésű, ami azt jelenti, hogy a program több megosztott, statikus könyvtárat tartalmaz.⁶

Az *OpenCV* tartalmaz képfeldolgozó modult, amely lineáris és nemlineáris képszűrést, geometriai képátalakításokat (átméretezés, affine és perspektivikus hajlítás, általános táblázatalapú újraleképezés), színtér-átalakítást végez. A videóelemző modul, ami mozgásbecslést, háttérkivonást és objektumkövetési algoritmusokat tartalmaz.

A mintaillesztés-alapú objektumfelismerési módszer az *OpenCV* implementációjában, a *matchTemplate* nevű függvényben valósul meg, amelynek paramétere a bemeneti forráskép, a bemeneti mintakép, a kimeneti kép és az illeszkedés mértékének meghatározására felhasznált módszer azonosítója.⁷ A bemeneti képeket 8 vagy 32 bites képek formájában várja a program, azonban fontos, hogy a két kép ugyanolyan formátumú legyen. A program az eredményeket az eredménymátrixban 32 bites lebegőpontos számokként tárolja el, függetlenül attól, hogy egycsatornás vagy színes (többcsatornás) képeket használtunk fel bemenetként. Az eredménymátrixból a legjobb találat kiválasztása a *minMaxLoc* nevű függvénnyel végezhető el.

Az illeszkedések mértékének meghatározására az *OpenCV* három alapvetően különböző módszert használ. A *négyzetes különbség (square difference)* módszere a hasonlóságot a két kép közti négyzetes különbségek alapján számítja ki. Ebben a módszerben a nulla jelenti a legnagyobb hasonlóságot, míg a legnagyobb szám a legkisebb hasonlóságot. A *korrelációs illesztés (correlation matching)* módszere a két képet összeszorozza, így a nagy értékek jelzik a jó hasonlóságot, a nulla pedig a kisebb hasonlóságot. A *korrelációegyüttható-illesztés (correlation coefficient matching)* módszere a képek átlagait felelteti meg egymással, azaz az 1 jelzi a jó hasonlóságot, a nulla jelzi, ha nincs hasonlóság, a -1 pedig a rossz hasonlóságot jelzi.

Ahhoz, hogy a módszert katonai rendfokozati jelzések felismerésére használjuk fontos, hogy a felismerendő jelzés jól látszódjon (megfelelő méretben, nézetben,

⁵ *OpenCV nyílt forráskódú számítógépes látástechnológia* [é. n.].

⁶ BRAHMBHATT 2013.

⁷ PAJANKAR 2015.

orientációban, megvilágításban) a kamera képén. Azt, hogy a felismerendő objektum szerepel-e a kameraképen, az eredménymátrixból a legnagyobb illeszkedésre utaló érték nagysága és mátrixban lévő pozíciója alapján becsülhetjük meg. Több felismerendő rangjelzés esetén több mintaképet használhatunk, amelyeket egyenként kereshetünk a kamerából kapott képen.

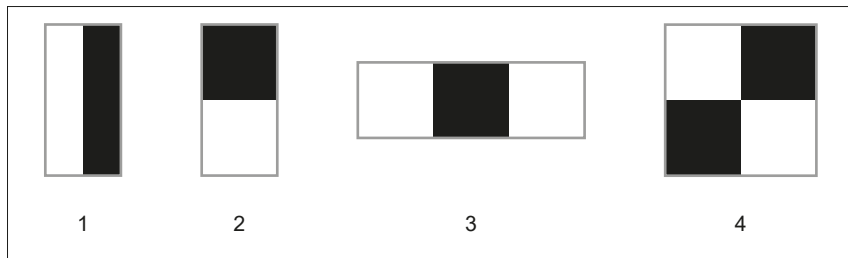
A kaszkádosított osztályozó alapú objektumfelismerés

A kaszkádosított osztályozó (*Cascade Classifier*) alapú felismerés manapság az egyik leghatékonyabb és legerősebb módszer általános objektumok való idejű detektálására. Jellemzője, hogy rövid feldolgozási idő alatt képessé válhat nehezen leírható, összetett, komplex objektumok felismerésére hosszabb tanítási folyamat után. Ezek az objektumok szinte tetszőlegesek lehetnek, például arcok, emberek, épületek vagy éppen katonai rendfokozati jelzések, amelyeknél a felismerés megfelelő tanítás esetén változó körülmények között is eredményes maradhat.

A kaszkádosított osztályozó alapú objektumfelismerés folyamata

A felismerés konkrét folyamata több jellegzetességtípus segítségével történik, ilyenek például a Haar-szerű jellegzetességek (*Haar-Like Features*), a *Local Binary Patterns (LBP)* jellegzetességek, illetve a *Histogram of Oriented Gradients (HOG)* jellegzetességek. Ezekből az *OpenCV* eszközei elsősorban a Haar-szerű jellegzetességeket és a LBP-jellegzetességeket támogatják.

A Haar-szerű jellegzetességek nevüket a *Haar Waveletek*hez való hasonlóságuk alapján kapták. A *Haar Wavelet* előnyös lehet a hirtelen átmenetekkel rendelkező jelek elemzésénél, például a gépek szerszám-meghibásodásának felügyeletéhez. Hirtelen átmenetekkel rendelkező jelek egy fényképen is lehetnek, például egy tárgy éleinél vagy derékszögnél, ahol a környező képkockák más tulajdonságot vesznek fel, mint a vizsgált képkocka. Egy Haar-szerű jellegzetesség szomszédos, azonos méretű és alakú téglalap alakú területeket vizsgál, az ilyen területeken kiszámítja a pixelintenzitások összegét, és ezeknek az összegeknek a különbségeit vizsgálja (egymáshoz képest). Ezen különbségek alapján az egyes területeket (egymáshoz képest) sötétnek vagy világosnak nyilvánítja, ami alapján a kép részeit kategorizálni tudja.



7. ábra. Haar-jellegzetesség-típusok

Forrás: a szerző szerkesztése

A 7. ábra bemutatja a Haar-szerű jellegzetességtípusokat, amelyek a következők: az úgynevezett kettőtéglalap- (1 és 2), háromtéglalap- (3) és négytéglalap- (4) jellegzetesség. Az arcon lévő szemnél például az alsó részén lehet egy 2-es Haar-jellegzetesség. A felső sötét kocka a szem, míg az alsó világosabb kocka a bőr. Így meg lehet különböztetni például a szemet az arcon, hiszen a szem körvonala élesen eltér a körülötte lévő bőr jellegzetességétől, azaz színétől.

A módszer eredeti változatát Paul Viola és Michael Jones publikálta 2001-ben.⁸ A módszer kifejlesztését elsősorban az arcfelismerés motiválta, amellyel sikerült az első valós idejű arcfelismerők egyikét megvalósítani, amely egy Intel Pentium III számítógépen 15 képkocka/másodperc sebességgel futott.

Az objektumfelismerési keretrendszer a Haar-szerű jellegzetességek felhasználásával képes elérni a gyors és hatékony felismerést. Ehhez szükséges még egy új képbábrázolás (integrált kép), amely segítségével lehetővé válik a felismerő által felhasznált jellegzetességek rendkívül gyors kiértékelése. Szükséges még egy osztályozó létrehozására felhasznált módszer, amely során *Adaptive Boosting* (*AdaBoost*) segítségével vizuális jellegzetességeket választunk ki nagyobb készletből. A harmadik lépésben szükséges egy módszer, amely lehetővé teszi a kép háttérterületeinek gyors kiszűrését, ezáltal több számítást felhasználva az ígéretesnek tűnő objektumszerű képrészletre.

Rainer Lienhart és Jochen Maydt a Haar-szerű jellegzetességek kibővített változatát publikálták 2002-ben.⁹ Ebben a változatban az objektumfelismerést olyan módon próbálták meg fejleszteni, amelyben bevezették a jellegzetességek 45°-kal elforgatott változatát. Ennek a kibővített változatnak a felhasználásával sok

⁸ VIOLA–JONES 2001.

⁹ LIENHART–MAYDT 2002.

helyzetben pontosabban leírhatók a felismerendő objektumok, így meglehetősen hatékonynak bizonyulnak. Az elforgatott jellegzetességek hatékony kiszámításához az integrálképek eredeti formájukban nem voltak felhasználhatók, ezért ennek egy módosított változatát definiálták. Az integrálkép elforgatott változatában (*Rotated Integral Image; RII*) a mezők értéke az eredeti, tőlük bal felső régióba eső pixelek helyett annak a régiónak az óramutató járásával ellentétesen 45°-kal elforgatott verziójába eső pixeleket összegzi.

Ezt az elforgatott integrálképet az eredeti kép pixelein kétszer végighaladva lehet kiszámítani, tehát az eredeti integrálképhez képest körülbelül kétszer lassabban számolható, ez azonban még mindig nagyon gyorsnak mondható.

A felismerés során adott méretű célablak mozog a bemeneti képen, és a Haar-szerű jellegzetességeket az ablak minden helyén kiszámítják. Annak érdekében, hogy több méretben is felismerhessük az objektumot, az osztályozó könnyen átméretezhető.

Ahhoz tehát, hogy egy ismeretlen méretű katonai rendfokozati jelzést is felismerjen a folyamat, a pásztázást iteratívan többször el kell végezni az osztályozó különböző méretű változataival. Ha a kaszkádban bármely osztályozó elveti a részterületet, a feldolgozás véget ér azzal a következtetéssel, hogy a katonai jelzés nem szerepel a képen. Ha a részterület sikeresen keresztüljut minden osztályozón, a katonai rendfokozati jelzés felismerése sikeresnek minősülhet.

Az OpenCV által támogatott kaszkádosított osztályozó alapú objektumfelismerés

Az OpenCV-ben a kaszkádosztályozók alapján való felismerés elvégzéséhez kapcsolódó funkciók a *CascadeClassifier* nevű osztályban találhatók meg. A *load* függvénnyel tölthetjük be az *OpenCV* eszközeivel korábban betanított XML formátumú kaszkádosztályozót, ez lehet Haar-szerű vagy LBP-jellegzetesség-típusú is.¹⁰ Ezután a detektálást a *detectMultiScale* függvénnyel végezhetjük el.

A függvény paramétere a bemeneti kép, a felismert objektumok helyét tartalmazó téglalapok listája (kimeneti paraméter), a skálázási faktor, a szomszédok minimális száma, egy *flags* mező, továbbá a felismerni kívánt objektum minimális és maximális mérete. Az objektumot a megadott minimális és maximális méret között keresik, a keresett objektum mérete minden iterációban a skálázási

¹⁰ BUENO GARCÍA et al. 2015.

faktorral változik meg. A nagyobb érték itt nagyobb megbízhatóságot, de kisebb pontosságot jelent.

Osztályozók létrehozására az *OpenCV* az *Opencv Traincascade* nevű eszközzel rendelkezik, amely a bemeneti képek és a megadott paraméterek alapján el tudja végezni egy kaszkádosztályozó betanítását.

Ahhoz, hogy a módszert katonai rendfokozati jelzések felismerésére használjuk, fontos a rendszer betanítása. Mivel ezek a jelzések (magyar és külföldi) száma elérheti az 1000-et, így a betanítási folyamat hosszúnak tűnhet, de az eredmény gyors felismerést tesz lehetővé.

A jellegzetességmegfeleltetés-alapú objektumfelismerés

A jellegzetességmegfeleltetés (*Feature Matching*) módszerével a képek jellegzetességeit közvetlenül is felhasználhatjuk a katonai rangjelzések felismerésére. A jellegzetességmegfeleltetés módszere jellegzetes pontokat keres a bemeneti képeken, majd az egymáshoz hasonló jellegzetes pontokat megpróbálja megfeleltetni egymásnak. Ha kellően sok pontot sikerült megfeleltetni a képek között egymásnak, és a pontok elhelyezkedése is konzisztens, az arra utalhat, hogy a képeken ugyanaz a katonai rangjelzés szerepel.

A jellegzetességmegfeleltetés-alapú objektumfelismerés folyamata

A jellegzetességek a képeken bizonyos szerkezetet, alakokat jelölnek, így ezek lehetnek egyszerűbb dolgok, mint például pontok vagy élek, de akár összetettebb szerkezetek is.

Adott jellegzetesség konkrét ábrázolása számos különböző módon történhet, például a kép egy pontján van-e él, vagy nincs, ezt jelölhetjük egyszerű logikai értékkel. Egy másik megközelítés szerint leírhatjuk a jellegzetességet a pont körüli terület átlagos színe vagy a pont kontrasztja alapján.

Egy jellegzetesség konkrét reprezentációját a jellegzetesség leírójának nevezzük. Időnként nem elegendő csak egyfajta jellegzetességeleíró ahhoz, hogy megfelelő információnk legyen a képről, így több jellegzetességeleíró is tartozhat a kép egy-egy pontjához. Az így kapott információkat gyakran egyetlen vektorba szervezik, amelyre jellegzetességvektor néven hivatkoznak.

Egy leíró lehet alacsonyabb szintű (amely például csak egyszerű élek jelenlétére utal) vagy összetettebb (amely például a pont nagyobb környezetét is pontosan jellemzi színek, alakok alapján vagy több egyéb szempontból).¹¹

A katonai rendfokozati jelzések felismeréséhez elegendő az élek felismertetése, hiszen láttuk, a 6. táblázat szerint az altisztek és a főtiszték közötti különbség egy csík a jelzésben, amelynek a felismertetése az élek felismertetésén alapszik. Mivel ez egy egyszerűbb leíróval lehetséges, így a felismerés is könnyebb és gyorsabb lesz, mivel az alacsonyabb szintű leírókat általában könnyű kiszámítani és ábrázolni is.

Az OpenCV által támogatott jellegzetességmegfeleltetés- alapú objektumfelismerés

Az *OpenCV*-ben a leírók többdimenziós tér vektoraiként vannak definiálva.¹² A könyvtár két főbb típusú leírót tartalmaz: a lebegőpontosat és a binárisat.

Az *OpenCV* közös interfészeket biztosít a folyamat minden fontos lépéséhez: a *FeatureDetector* a jellegzetesség detektálásához, a *DescriptorExtractor* a jellegzetesség kivonatolásához és a *DescriptorMatcher* a jellegzetesség megfeleltetéséhez.

A *FeatureDetector* interfész elvégzi a jellegzetességek keresését a megadott képen vagy képeken, és kimenetként *KeyPoint* típusú adatokat tartalmazó halmazt hoz létre. A halmaz tartalmazza többek között a pontok koordinátáinak, méretének, orientációjának és illeszkedése mértékének adatait.

A *DescriptorExtractor* interfész funkciója a *detect* metódus, amely a *KeyPoint* típusokként megadott kulcspontok helyein kiszámítja a jellegzetességek leíróit, azaz létrehoz egy mátrixot, amelynek minden sora egy kulcsponthoz tartozó leíró vektor.

A *DescriptorMatcher* interfész a leírók egyik halmazában lévő elemekhez megkeresi a másik halmazból a legközelebbi leírókat. Az eredményt az interfész *DMatch* adattípusként adja vissza, amely tartalmazza a találatokhoz kapcsolódó két leíró azonosítóját, a kép azonosítóját és a távolságot, amely a találat minőségére utal.

¹¹ SHOJANIA 2003.

¹² JOSHI – MILLAN ESCRIVA – GODOY 2016.

A módszer segítségével megvalósított katonai rendfokozatijelzés-felismerés hatékonyságát és sebességét döntő mértékben meghatározzák a folyamat egyes fázisai (detektálás, leírókészítés, megfeleltetés) során felhasznált konkrét módszerek tulajdonságai és a folyamat kimenetének (egymásnak megfeleltetett leírópárok) konkrét felhasználási módja is. Mivel a módszer leginkább a fixen sok jellegzetes, jól leírható ponttal rendelkező objektumok esetén működik kiválóan, így a katonai rendfokozati jelzés felismerésnél is használható.

Megfelelő módszerek felhasználásával a felismerés futási sebessége kielégítő lehet okostelefonokon való közel valós idejű felhasználáshoz is. A tárhelyigény a felismerendő katonai rendfokozati jelzéseket tartalmazó adathalmaztól függ, azok mennyiségétől és méretétől.

Következtetések

Az MI az évek múlásával egyre komplexebb feladatok ellátására lesz alkalmas. A távoli jövőben elképzelhető, hogy képes lesz az emberi gondolkodást utánozni, azaz helyettesíteni az embert, a katonát. Az MI alkalmas átvenni számos katonai számítási és következtetési feladatot, kiterjeszti az emberi gondolkodást, így jobb döntést hozhat, mint az ember.

Számtalan robotokkal és MI-vel kapcsolatos cikk, írás jelenik meg nap mint nap a világhálón és a tudományos folyóiratokban. A legtöbbjük az MI-ről készült anyagokban az MI felhasználási területét, az ember mindennapjaira gyakorolt hatását és veszélyeit emeli ki. Ez a publikáció rávilágít egy fontos katonai alkalmazási területre, ahol az MI-vel támogatott gépi látást használhatjuk a katona látás útján történő objektumfelismerési képességének növelésére.

A gépi látás központi eleme egy neurális háló, amely feldolgozza a kapott értékeket, s kiszámítja a megoldást. A neurális hálózatok elterjedése magával hozta az azokat támogató, nyílt forráskódú modellező rendszerek nagy számát, mint például a publikációban tárgyalt *OpenCV*. A nyílt forráskód lehetővé teszi a katonai alkalmazást is, így a publikációban említett katonai rendfokozati jelzések felismerését is elérhetővé teszi. A legeredményesebb módszer talán a jellegzetességmegfeleltetés-alapú objektumfelismerés lehet a katonai rendfokozati jelzések felismertetésére. Ez a felismertetés gyors, és akár a szolgálati mobiltelefonon is lehetne futtatni a programot.

A dolgozatban bemutatott felismertetési módszerek sebességének mindegyike függ a felismertetendő objektumok számától. A betanítási folyamat pedig

a felismertetendő objektumokról készült fényképek darabszámától is függhet, és akár több napot is igénybe vehet.

Fontos felhasználási terület lehet még a katonai műveleti területen feladatot végrehajtó digitális katona támogatása olyan gépi látással működő eszközzel, amely esetleg integrálható a rendszerébe. A gépi látás képes működni, amikor a katona alszik, vagy képes látni a katona helyett, és őrizni a biztonságát. A modern harcokban már drónokat is használnak minden katonai és civil célpont ellen, így a digitális katona ellen is. A drónok gyorsan, a legváratlanabb időben csapnak le áldozatukra, így ezek időbeni felismerése kritikus pontja a lövészárokban szolgálatot teljesítő katonák túlélésének.

Felhasznált irodalom

2012. évi CCV. törvény a honvédek jogállásáról. 1. melléklet: Az állománycsoportok, a rendfokozati csoportok és a rendfokozatok.
- BRAHMBHATT, Samarth (2013): *Practical OpenCV*. New York: APress. Online: <https://doi.org/10.1007/978-1-4302-6080-6>
- BUENO GARCÍA, Gloria et al. (2015): *Learning Image Processing with OpenCV*. Birmingham: Packt.
- JOSHI, Prateek – MILLAN ESCRIVA, David – GODOY, Vinicius (2016): *OpenCV by Example*. Birmingham: Packt.
- KOLLÁR Csaba – NAGY Barna (2021): A mesterséges intelligencia felhasználási lehetőségei az objektumfelismerésben. *Biztonságtudományi Szemle*, 3(2), 123–140.
- LIENHART, Rainer – MAYDT, Jochen (2002): An Extended Set of Haar-Like Features for Rapid Object Detection. *IEEE International Conference*. Rochester, 22–25 September.
- OpenCV nyílt forráskódú számítógépes látástechnológia* [é. n.]. Online: <https://docs.opencv.org/3.2.0/>
- PAJANKAR, Ashwin (2015): *Raspberry Pi Computer Vision Programming*. Birmingham: Packt.
- PORKOLÁB Imre – NÉGYESI Imre (2019): A mesterséges intelligencia alkalmazási lehetőségeinek kutatása a haderőben. *Honvédségi Szemle*, 147(5), 3–19. Online: <https://honvedelem.hu/images/media/5f2bd1646eeb8298912683.pdf>
- SHOJANIA, Hassan (2003): *Real-Time Traffic Sign Detection*. Online: <https://hassan.shojania.com/pdf/TrafficSignDetection-Paper.pdf>
- VIOLA, Paul – JONES, Michael (2001): Rapid Object Detection Using a Boosted Cascade of Simple Features, in Computer Vision and Pattern Recognition. *Proceedings of the 2001 IEEE Computer Society Conference*, 1, I–511.