

Ádám Balázs¹ 

Mesterséges intelligencia a tűzszerészfeladatokban – Tűzszerész Támogató Információs Rendszer szoftveres alapjai, 3. rész²

Artificial Intelligence in Explosive Ordnance Disposal Tasks – Software Basics of EOD Support Information System, Part III

A „Mesterséges intelligencia a tűzszerészfeladatokban” című négyrészes cikksorozat fő célja a Mesterséges Intelligencia Alapú Tűzszerész Támogató Információs Rendszer bemutatása. A cikksorozat 3. részében szó esik a képelemzés szoftveres és matematikai hátterének lényegéről, ami alapján a mesterséges intelligencia egy adott képet az előre meghatározott csoportok valamelyikébe sorol. A képelemző szoftver működése jelenti a tűzszerészek támogatását végző program legfontosabb elemét.

Kulcsszavak: *tűzszerész, mesterséges intelligencia, robbanótestek, felismerő rendszer, aknagránát*

The main goal of the four-part article series entitled Artificial Intelligence in EOD Tasks is to present the EOD Support Information System Based on Artificial Intelligence. In Part 3 of the series of articles, the essence of the software and mathematical background of image processing, on the basis of which the artificial intelligence classifies a given image into one of the predefined groups. The operation of the image processing software is the most important element of the EOD supporting program.

Keywords: *EOD, artificial intelligence, explosive ordnance, recognition system, mortar rounds*

¹ Szakasparancsnok, MH 1. Tűzszerész és Folyamőr Ezred, e-mail: adam.balazs@mil.hu

² A publikáció az Innovációs és Technológiai Minisztérium ÚNKP-21-1-I-NKE-93 kódszámú Új Nemzeti Kiválóság Programjának a Nemzeti Kutatási, Fejlesztési és Innovációs Alapból finanszírozott szakmai támogatásával készült.

Bevezetés

A Mesterséges Intelligencia Alapú Tűzserész Támogató Információs Rendszer célja, hogy a tudomány jelen állása szerinti legkorszerűbb feltételeket nyújtsa a tűzserész műveletek irányításának és szervezésének támogatására. Ezt a vezetési támogatást a mesterségesintelligencia-alapú képelemző szoftver biztosítja, amely valós idejű és szűrt, tehát megfelelően hitelesített adatokat biztosít a parancsnok részére a minél pontosabb és körülményesebb döntéshozatalhoz. A rendszer gyakorlati felhasználására két területen, egyrészt a rendőrfőkapitány, másrészt a tűzserészfőkapitány tevékenysége során kerülhet sor.

Mesterségesintelligencia-alapú tűzserész-támogató információs rendszer

A mesterségesintelligencia-, vagyis a MI-³ alapú szoftvert a lakossági bejelentést követően a kiterjedő rendőrfőkapitány tudná első lépésben felhasználni, aki ennek segítségével meghatározhatja a robbanótest kategóriáját (nyolc alapvető kategóriát különböztettem meg), illetve a tűzserész műveleti vezetés kérésére további csoport és alcsoport, tehát a konkrét eszköz meghatározása is történhet. Ennek és az eszköz helyzetének ismeretében sokkal egyszerűbbé válna a robbanótest-mentesítés prioritásának (sürgősségének) meghatározása, valamint következtetni lehetne a mentesítéshez szükséges alapvető felszereléseken (például robbanóanyag) túl az esetleges speciális eszközre (például csörlő, emelődaru, vegyvédelmi biztosítás) is. Másik fontos alkalmazási területet jelentene a bejelentések során a tényleges robbanótestek és a nemrobbanó testek (NRT) megkülönböztetése. Ugyanis a bejelentések hozzávetőleg 10%-a nemrobbanó test,⁴ ami az éves átlag (2000 riasztás) esetében már nem elenyésző szám.⁵ A szoftver segítségével az ezekre a bejelentésekre fordított személyi, anyagi és technikai kiadásokat lehetne megspórolni. A fejlesztés egy későbbi szakaszát követően a tűzserészfőkapitány feladat-végrehajtását is támogathatja a rendszer, amennyiben a hosszú tanítási folyamat eredményeként a tévesztés kockázata már minimálisra redukálódott. Ebben az esetben a felismerő szoftver önállóan képes lenne azonosítani a robbanótestet, és egy tűzserész mentesítő robot vezérlésébe integrálva azt képes lehet a hatástalanítás egyes meghatározott folyamatait autonóm módon végrehajtani.

A támogatás lényege, hogy az előtalált robbanótestről a Magyar Honvédség (MH) Tűzserész Ügyeletére beérkező első információk alapján minél pontosabban behatárolják az eszköz kategóriáját és típusát. Az így megszerzett információk alapján az eszköz veszélyességi foka is sokkal nagyobb valószínűséggel behatárolható, illetve a feladatra küldhető, átirányítható főkapitány felszerelésének, fogyó anyagainak és robbanószer mennyiségének függvényében azok sokkal nagyobb precizitással, hatékonysággal irányíthatók át egyik feladatról a másikra. Ezek

³ Angol megnevezése: Artificial Intelligence – AI.

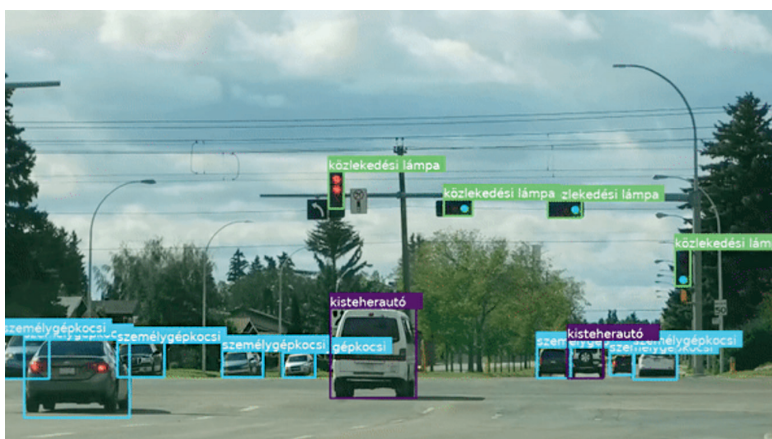
⁴ EMBER 2020a: 59–63.

⁵ EMBER 2020b: 32–42.

olyan képességek, amelyek akár már néhány éven belül is az MH 1. Tűzserész és Folyamőr Ezred (a továbbiakban: Tűzserész Ezred) parancsnoki állománya számára átfogó és hiteles döntéstámogatást nyújthatnak.

Objektumfelismerés vagy képelemzés

Napjaink MI-kutatásai legnagyobb számban az objektumfelismerés vagy képelemzés témájában zajlanak. Az objektumfelismerés kifejezést gyakran szinonimaként használják a képelemzés és képfeldolgozás szavakra, de a két fogalom között van lényeges eltérés is.⁶



1. ábra: Objektumfelismerés a valóságban

Forrás: Objektumfelismerő (képfelismerő) AI-megoldásunk 2019

Az objektumfelismerés fogalmát leggyakrabban azokra a szoftverekre kell érteni, amelyek élő képen, de inkább videós anyagon képesek kategorizálni és felismerni az előre betáplált alakzatokat. Ezek egymástól gyökeresen eltérő méretű, formájú, színű és egyéb tulajdonságú tárgyak, eszközök. Éppen ezért a program célja csupán a tárgyak egymástól történő megkülönböztetése (1. ábra) és nem az azonos tárgyak közötti különbségek vizsgálata.⁷

Ezzel szemben a képelemzés (Image Processing) lényege, hogy az egymáshoz nagyon hasonló tárgyak, tárgykategóriák között az adatbázis felhasználásával a tanulási folyamat eredményeként a létrejövő neurális háló képes legyen minimális, apró színbeli, formabéli vagy más eltéréseket azonosítani, majd betanulni, és így valamekkora bizonytalansággal a későbbiekben a felhasználás során meg is különböztetni azokat. Tehát a képelemzés már sokkal precízebb és pontosabb beazonosítást eredményez,⁸ mint az objektumfelismerés, de

⁶ Objektumfelismerő (képfelismerő) AI-megoldásunk 2019.

⁷ SHARMA 2019; KORSTANJE 2020.

⁸ A képelemzés két fajtáját különböztetik meg: az analóg és digitális felismerést. Analóg felismerés esetén egy elkészült kép egyes pontjaihoz, pixeleihez rendelnek színértéket, így a kép egy számokkal leírható adathalmazzá válik, ami utána már változtatható. Ezt a technológiát használják a régi képek vagy videók digitális újrászínzéséhez.

amennyiben a nagy eltérésekkel rendelkező objektumok valós idejű felismerése a cél, abban az esetben az objektumfelismerés eszközeinek alkalmazása a leghatékonyabb megoldás.⁹

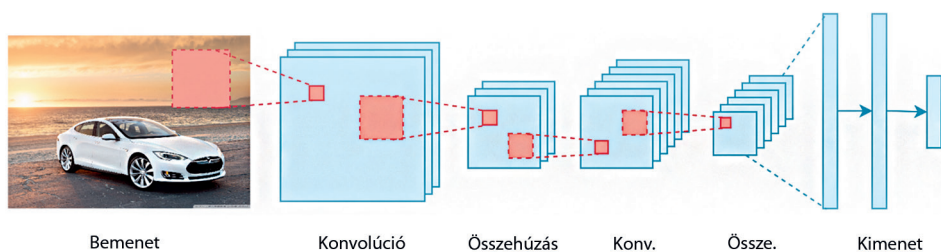
Az objektumfelismerés- és a képelemzés-technológiákat alkalmazó számos katonai és polgári fejlesztés látott már napvilágot, ezekről terjedelmi megfontolások miatt nem esik szó, de a hivatkozott cikkekben többet is meg lehet tudni róluk.¹⁰

Az objektumfelismerés tűzszerész szakfeladatokban történő egyik alkalmazási lehetősége lenne az improvizált robbanóeszközök¹¹ egyes elemeinek, mint például a robbanóanyag-töltet, a detonátor, az indító mechanizmus, az áramforrás és vezetékek, az időzítő berendezés, a hatásfokozó repeszek vagy a rejtést szolgáló különböző álcázó burkolat felismerése.¹² Ezek azonosítása kiemelt fontosságú, ugyanis a Közel-Kelet műveleti területein a saját csapatainkra mért támadások jelentős részét ilyen IED-k alkalmazásával hajtják végre.¹³

A képelemzés elvi működése

A konvolúciós neurális hálóról¹⁴ a cikksorozat 2. részében már esett szó, ezúttal pedig célszerű azok működési elvét is megvizsgálni (2. ábra), hogy megértsük szerepüket a képelemzés területén. A konvolúciós hálózatok jelentősen eltérnek a többi neurális háló-típusoktól. Ez leginkább azzal magyarázható, hogy ezeket a hálókat kifejezetten képi felismerésre fejlesztettek ki.¹⁵

Fő előnye elődeihez képest, hogy emberi felügyelet nélkül, automatikusan is képes felismerni a fontosabb eltéréseket. Például macskákról és kutyákról készült képek alapján képes megtanulni az egyes csoportok jellegzetes vonásait, így azokat meg tudja különböztetni egymástól. Minden CNN-modell hasonló architektúrát követ, amint az az 2. ábrán látható.¹⁶



2. ábra: Konvolúciós folyamatok

Forrás: DERTAT 2017 alapján a szerző szerkesztése

A digitális képek esetében pedig a képek pixeleihez már alapvetően hozzá van rendelve egy adat (x, y koordináta és szín), így az átalakítás nem szükséges. *Analog Image Processing vs. Digital Image Processing* [é. n.].

⁹ *Image Processing Definition, Examples and Application* [é. n.].

¹⁰ NÉMETH-VIRÁGH 2022: 2–7; NÉMETH-VIRÁGH 2023: 2–6.

¹¹ Angol megnevezése: Improvised Explosive Device – IED.

¹² KOVÁCS 2012: 37–52.

¹³ TOMOLYA-PADÁNYI 2012: 34–67.

¹⁴ Angol megnevezése: Convolutional Neural Network – CNN.

¹⁵ TÓTH 2019.

¹⁶ *What Is Deep Learning? 3 Things You Need to Know* [é. n.].

Az elemzésre váró kép a rendszer bemeneti információja, ez az input. A háló a kép egyes kijelölt részeiről egy számot generál a konvolúció matematikai formuláival. A konvolúciós műveletek összekapcsolódásával, azok sorozataként jön létre a neurális háló, amelyet számos teljesen összekapcsolt réteg alkot. A folyamat végén, az osztályozás (többosztályos) során létrejön a kimenet, a softmax. A CNN fő eleme a konvolúciós réteg. A konvolúció matematikai művelet két információhalmaz összevonására. Esetünkben a konvolúciót a bemeneti adatokra alkalmazzuk egy konvolúciós szűrő segítségével, így létrejön a jellemzőtérkép.¹⁷

1	1	1	0	0
0	1	1	1	0
0	0	1	1	1
0	0	1	1	0
0	1	1	0	0

3. ábra: Konvolúció bemenete

Forrás: DERTAT 2017 alapján a szerző szerkesztése

1	0	1
0	1	0
1	0	1

4. ábra: Konvolúciós szűrő

Forrás: DERTAT 2017 alapján a szerző szerkesztése

A példa kedvéért a 3. ábrán látható konvolúciós réteg a bemeneti kép, míg a 4. ábrán a konvolúciós szűrő, az úgynevezett kernel. A szűrő alakja miatt a példában bemutatott folyamatot 3×3-as konvolúciónak nevezik. A konvolúciós műveletet úgy hajtjuk végre, hogy ezt a kernelt a bemenetre helyezzük. Minden helyen elemenkénti mátrixszorzást végzünk a kernelben és a bemenetben található értékekre, majd az eredményt összegezzük. Ez az összeg kerül be a jellemzőtérképbe. Azt a „zöld területet”, ahol a konvolúciós művelet végbemegy, receptív mezőnek nevezzük. A szűrő méretéből adódóan a receptív mező is 3×3-as méretű. Az 5. ábrán látható, ahogy a szűrő a bal felső sarokban van, a mátrixműveletek elvégzése után pedig, a konvolúciós művelet kimeneteként a jellemzőtérképen megjelenik a „4” kapott érték (6. ábra). Ezután jobbra kell csúsztatni a kernelt, és végre kell hajtani ugyanazt a műveletet, hozzáadva az eredményt a jellemzőtérképhez. Ezt a módszert folytatva elkészül a teljes térkép (7. ábra).¹⁸

¹⁷ DERTAT 2017.

¹⁸ DERTAT 2017; BROWNLEE 2020.

1×1	1×0	1×1	0	0
0×0	1×1	1×0	1	0
0×1	0×0	1×1	1	1
0	0	1	1	0
0	1	1	0	0

5. ábra: Bemenet és a kernel

Forrás: DERTAT 2017 alapján a szerző szerkesztése

4		

6. ábra: Jellemzőtérkép részlet

Forrás: DERTAT 2017 alapján a szerző szerkesztése

4	3	4
2	4	3
2	3	4

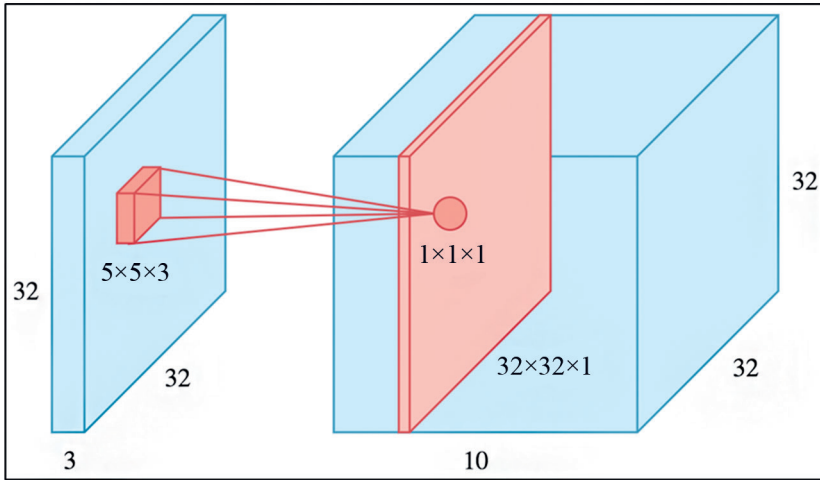
7. ábra: Teljes jellemzőtérkép

Forrás: DERTAT 2017 alapján a szerző szerkesztése

A példa szerinti konvolúciós műveletet két dimenzióban (2D) mutattuk be, 3×3-as szűrő használatával, de a tényleges felhasználás során ezeket a konvolúciókat három dimenzióban (3D) hajtják végre. A valóságban egy kép 3D-s mátrixként jelenik meg magassággal (y koordináta), szélességgel (x koordináta) és mélységgel, ahol a mélység a színcsatornáknak (RGB)¹⁹ felel meg. A konvolúciós szűrőnek meghatározott magassága és szélessége van, például 3×3 vagy 5×5, és kialakításánál fogva lefedi a bemenet teljes mélységét, ezért is kell 3D-snek lennie. Az eredményes vizsgálathoz több konvolúciót kell végrehajtani egy bemeneten, mindegyik esetben más-más szűrőt használva, amely így eltérő jellemzőtérképet eredményez. Ezt köve-

¹⁹ Red–Green–Blue színmodell.

tően egyesíteni kell ezeket a jellemzőtérképeket, és ez lesz a konvolúciós réteg végső kimenete. A 8. és 9. ábrákon egyetlen szűrő alkalmazásával készült folyamatra láthatunk példát.²⁰

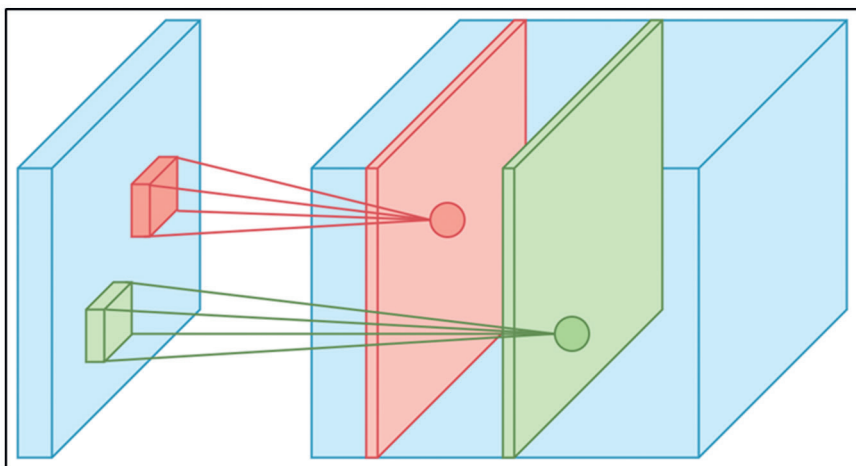


8. ábra: 3D konvolúció
 Forrás: DERTAT 2017

Tegyük fel, hogy van egy $32 \times 32 \times 3$ -as képünk, és egy $5 \times 5 \times 3$ -as kernelt használunk, azaz a konvolúciós szűrő mélysége megegyezik a kép mélységével. A szűrő egy adott helyen lefedi a bemenet egy kis részét, itt végrehajtja a fent leírt konvolúciós műveletet. Az egyetlen különbség az, hogy ezúttal a mátrixszorzás összegét 3D-ben végzi, és nem 2D-ben, de az eredmény továbbra is műveletenként egy-egy skalár. A bemeneten a szűrő áthelyezésével (a fenti módon) minden helyen megtörténik a konvolúció, amely eredményét a jellemzőtérkép összesíti. Ez a jellemzőtérkép $32 \times 32 \times 1$ méretű, a 8. ábra jobb oldalán piros szeletként látható. Ha 10 különböző szűrőt használunk, akkor 10 db $32 \times 32 \times 1$ méretű tereptérképet kapnánk, és ezeket a mélységi dimenzió mentén egymásra rakva megkapnánk a konvolúciós réteg végső kimenetét: egy $32 \times 32 \times 10$ méretű 3D mátrixot, amely a 8. ábra jobb oldalán nagy kék dobozként látható. A tereptérkép magassága és szélessége változatlan, azaz továbbra is 32, ami a kitöltésnek köszönhető. A 9. ábrán az látható, hogyan készül két jellemzőtérkép a mélységdimenzióban eltolva egymástól. Az egyes szűrők konvolúciós műveleteit egymástól függetlenül hajtják végre, és az eredményül kapott jellemzőtérképek diszjunktak, vagyis nincs közös részük.²¹

²⁰ DERTAT 2017.

²¹ DERTAT 2017.



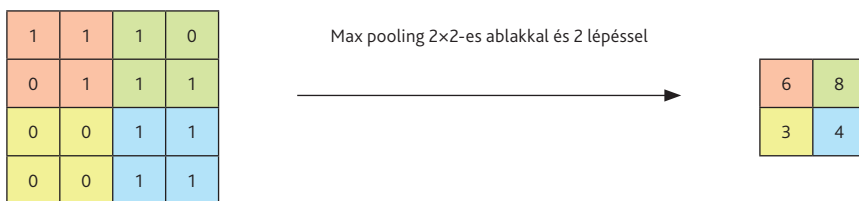
9. ábra: 3D konvolúció mélységi eltolással

Forrás: DERTAT 2017

A konvolúció után általában pooling (összehúzás) művelet történik a magasság és a szélesség csökkentése érdekében. Ez lehetővé teszi a paraméterek számának csökkentését, ami egyrészt lerövidíti a tanulási folyamat idejét, másrészt pedig a túltanulás kockázatát csökkenti. Az egyesített rétegek külön-külön leképezik az egyes tereptárgyak jellemtérképét, csökkentve a magasságot és a szélességet, a mélységen viszont nem változtatnak. A leggyakrabban használt összehúzó eljárás a max pooling, amely azt jelenti, hogy a pooling ablakban található maximális értéket veszi fel az adott terület értékének. A konvolúciós művelettel ellentétben a poolingnek nincsenek paraméterei. Egy ablakot csúsztat a bemenete fölé, és egyszerűen felveszi az ablakban található értékek közül a maximális értéket. A konvolúcióhoz hasonlóan meg kell határozni az ablak méretét és lépésszámát. Példaként egy 2×2 -es ablak és 2-es lépésszám²² összevonás eredménye látható a 10. ábrán. Minden szín más ablakot jelöl. Mivel mind az ablak mérete, mind a lépésszám 2, az ablakok nem fedik egymást. Ennek a folyamatnak a célja, hogy csökkenjen a jellemtérkép szélessége és magassága. Összehasonlítva a jellemtérkép méreteit az összevonás előtt és után azt kapjuk, hogy ha a gyújtóréteg bemenetének mérete például $32 \times 32 \times 10$ volt, akkor a pooling paraméterek használatával az eredmény egy $16 \times 16 \times 10$ -es jellemzőtérkép lesz. A térkép magassága és szélessége is a felére csökken, de a mélység nem változik, mert a pooling egymástól függetlenül működik a bemenet minden egyes mélységi szeletén. A bemeneti kép 2D koordinátái (súlyok) száma negyedére csökkent. Figyelembevéve, hogy a konvolúciós hálók általában több millió súllyal dolgoznak, nagyon fontos az adattömeg csökkentése. A CNN-architektúrákban a pooling rendszerint 2×2 ablakkal és 2 lépéssel, míg a konvolúció 3×3 -as ablakkal, 1 lépéssel történik.²³

²² Az ablak kettő egységet mozog oldalra. Ezt az értéket konvolúció során érdemes 1-nek hagyni, pooling során pedig többnek is lehet, amennyiben megegyezik az ablak méretével, nem lesz átfedés az ablakok között.

²³ DERTAT 2017; BROWNLIE 2020; *Max Pooling* [é. n.].



10. ábra: Max pooling

Forrás: DERTAT 2017 alapján a szerző szerkesztése

A teljes felismerési folyamat nagyrészt ezekből az alapfolyamatokból áll, kiegészülve még olyan tevékenységekkel, amelyekről terjedelmi megfontolás miatt a publikációban ezúttal nem esik szó, de lényegében ezeknek az értékeknek az összehasonlításán alapszik a képi felismerés elve. Tehát a mintahalmaz ilyen módon létrejövő értékei kerülnek összehasonlításra a bemeneti kép értékével, ami alapján a mesterséges intelligencia meghatározza, hogy a kép melyik csoportba tartozik.

Szoftveres háttér

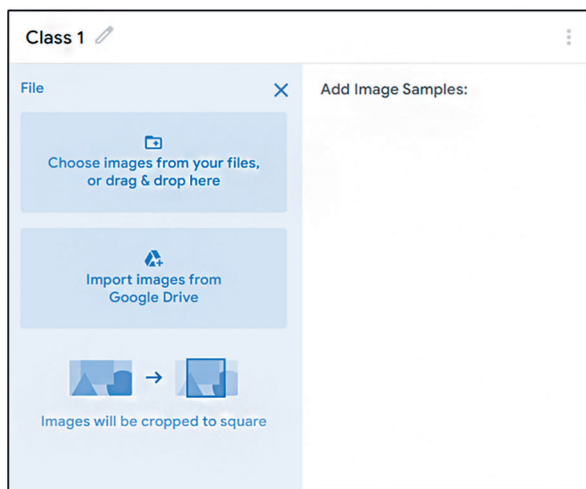
A kutatás kezdetén fontos volt egy olyan szoftver keresése, amely megfelelő hatásfokkal és pontossággal dolgozik, mégis kevés programozási ismeretre van szükség használatához. Az MI Koalíció *Bevezetés a Mesterséges Intelligencia világába* című kurzusát hallgatva ismertem meg a „Teachable Machine” alkalmazást, amelyet a kutatás során is használtam.

A Teachable Machine egy webalapú eszköz, a Google fejlesztése, amely gyors és egyszerű használatot biztosít, így mindenki számára elérhetővé teszi a gépi tanulási modellek létrehozását akár kép, zene vagy testhelyzetek felismerésére van szükség. A webes alkalmazás első verzióját 2017-ben ismerhették meg a felhasználók. Azóta számos fejlesztés és javítás történt nemcsak a grafikus kezelőfelületen, hanem a neurális hálókészítés beállításainak lehetőségeiben, valamint a hálóképzés sebességében és pontosságában is. Segítségével képfelismerő algoritmus készíthető a készülék kamerája vagy a meglévő fényképek használatával. Képes hangok és dallamok felismerésére is, mikrofon vagy hangfájlok alkalmazásával, továbbá képessé tehető az emberi mozdulatok és testtartások felismerésére valós időben webkamera segítségével.²⁴

A neurális háló képzése három egyszerű lépésben történik, amelyek a következők: adatgyűjtés és osztályozás, modellképzés és az elkészült modell exportálása, felhasználása. Adatgyűjtés során egy megfelelő mennyiségű és minőségű adattal feltöltött halmazt kell létrehozni (ezek a kívánt tevékenység alapján állhatnak kép, hang vagy videófájlokból). A feltöltött elemeket csoportosítani szükséges, ami a példák osztályokba vagy kategóriákba rendezését jelenti, ezeket kell majd a neurális hálónak megtanulnia. Az adatbázis feltöltése történhet

²⁴ Teachable Machine [é. n.].

a webkamerával valós időben készített képekből, a Google Drive-ról,²⁵ illetve adattároló eszközökről png-, jpeg- vagy bmp-kiterjesztéssel (11. ábra). A képeket a program automatikusan négyzet alakúra vágja ki. Kétféle feltöltési lehetőség közül lehet választani, 240 × 240 pixel méretű színes képek vagy 96 × 96 pixel méretű fekete-fehér képek. Természetesen a jobb képminőség és színes adathalmaz jobb felismerhetőséget is eredményez, így ajánlott ezzel a módszerrel használni az alkalmazást.²⁶



11. ábra: Adatok osztályokba rendezése

Forrás: Teachable Machine [é. n.]

Modellképzés során (12. ábra) a tanítás gombra kattintva, megkezdődik a tanulási folyamat, amelyhez a rendszer nem a saját számítógép erőforrásait használja fel, hanem a Google szervereit, így minimalizálja a felhasználó hardverigényét és maximalizálja a feladat elvégzésének sebességét. A tanítás során két változó értékét lehet állítani: az epoch²⁷ és batch size²⁸ változókat. A rövid időn belül elkészülő modell azonnal tesztelhetővé válik, így hamar megállapítható, hogy a modell képes-e megfelelően besorolni az új teszt példákat. A pontosság vizsgálatához próbafájlokat lehet készíteni a webkamera segítségével, vagy azok feltölthetők adattárolóról is. Az alkalmazás a tesztfájltra vetítve meghatározza, hogy az egyes csoportok, osztályokkal milyen százalékos arányban egyezik meg a tartalom. A modell exportálása,

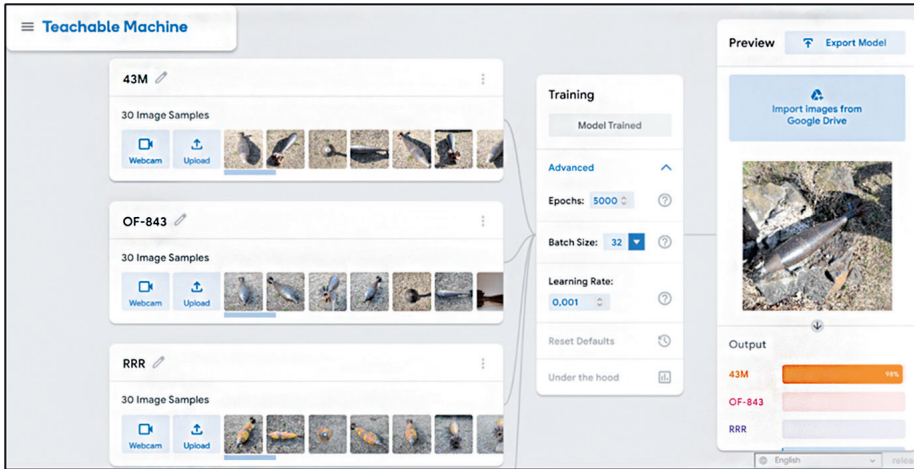
²⁵ A Google Drive egy fájlok tárolását, szinkronizálását és szerkesztését lehetővé tevő szolgáltatás, amelyhez minden Google-fiókkal rendelkező felhasználó számára 15 GB felhőalapú tárhelyet biztosít a cég.

²⁶ Teachable Machine [é. n.].

²⁷ Egy epoch a mintahalmaz egyszeri végigfuttatása a neurális hálón. Ha például ez az érték 50, akkor 50 esetben fut végig a minta feldolgozása és a súlyozás. A kis epochszám alultanított, a nagy epochszám túltanított neurális hálót eredményez. De alapvetés, hogy minél nagyobb az epochszám, várhatóan annál jobb modellt kapunk.

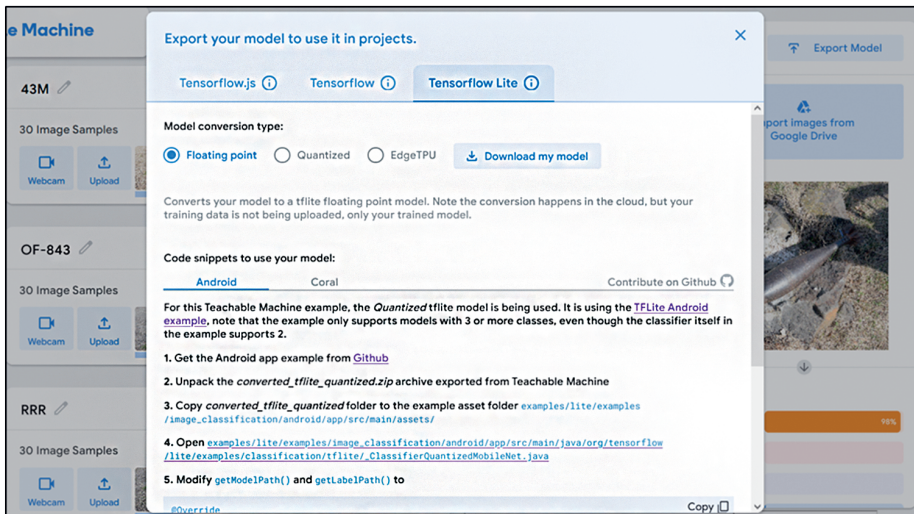
²⁸ A batch size lényege, hogy egy tanulási ciklusban meghatározott számú képet tanítunk be a neurális hálóba. Ha a tanulóhalmaz 80 képet tartalmaz és a batch size 16, akkor $80 \div 16 = 5$ batch-be lesz szétosztva a 80 kép. Ha mind az 5 batch átment a tanulás folyamatán, akkor kész egy epoch. Ha az 5 batch kétszer is átment a tanulási folyamaton, akkor kész kettő epochs és így tovább.

felhasználása során az elkészült neurális háló különböző formátumokban (Tensorflow.js, Tensorflow [Keras] és Tensorflow Lite) kimenthető az alkalmazásból (13. ábra), hogy azt a különböző projektekben lehessen alkalmazni. Egy ilyen plusz felhasználási lehetőséget biztosít a Google is, így az Android operációs rendszerekkel rendelkező telefonokra alkalmazást lehet programozni a kapott mesterséges intelligencia használatára.²⁹



12. ábra: Modellképzés

Forrás: Teachable Machine [é. n.]



13. ábra: Modell exportálása és letöltése

Forrás: Teachable Machine [é. n.]

²⁹ Teachable Machine [é. n.].

A program megfelelő működéséhez és használatához szükség van a már sokszor említett nagyméretű képi adatbázisra, amelynek kialakításához a Tűzserész Ezred laktanyájában, a kiképző kabinetben készítettünk fényképeket. Az általam használt adatbázisról a cikksorozat következő részében esik szó.

Összegzés

A cikkben bemutattam azokat a matematikai és informatikai megoldásokat, amelyek segítségével a képelemző mesterséges intelligencia működik. Ezek a folyamatok biztosítanak lehetőséget, hogy a program adott képet a korábban definiált csoportok valamelyikéhez sorolja. A kutatás során ehhez a feladathoz a Teachable Machine webalapú eszközt használtam, amely megfelelő határfokkal készített a kívánalmaknak megfelelő neurális hálókat. Az elkészült neurális hálók exportálhatók a programból, megkönnyítve ezzel a további fejlesztés és például egy telefonos applikáció létrehozásának lehetőségét.

Felhasznált irodalom

- Analog Image Processing vs. Digital Image Processing* [é. n.]. Online: www.javatpoint.com/analog-image-processing-vs-digital-image-processing
- BROWNLEE, Jason (2020): *How Do Convolutional Layers Work in Deep Learning Neural Networks?* Online: <https://machinelearningmastery.com/convolutional-layers-for-deep-learning-neural-networks/>
- DEKOR, Arden (2017): *Applied Deep Learning–Part 4: Convolutional Neural Networks*. Online: <https://towardsdatascience.com/applied-deep-learning-part-4-convolutional-neural-networks-584bc134c1e2>
- EMBER István (2020a): A lőszermentesítés szerepe az építőiparban. *Építőanyag*, 72(2), 59–63. Online: <https://doi.org/10.14382/epitoanyag-jsbcm.2020.9>
- EMBER István (2020b): The Role and the Risks of Explosive Ordnance Decontamination in Hungary. *Science & Military (Veda a Vojenstvo)*, 16(1), 32–42. Online: <https://doi.org/10.52651/sam.a.2021.1.32-42>
- Image Processing Definition, Examples and Application* [é. n.]. Online: www.zr-tech.co.uk/Image-Processing/
- KORSTANJE, Joos (2020): *What is the Difference Between Object Detection and Image Segmentation?* Online: <https://towardsdatascience.com/what-is-the-difference-between-object-detection-and-image-segmentation-ee746a935cc1>
- KOVÁCS Zoltán (2012): Az improvizált robbanóeszközök főbb típusai. *Műszaki Katonai Közlöny*, 22(2), 37–52. Online: https://mkk.uni-nke.hu/document/mkk-uni-nke-hu/2012_2_03%20IED-k%20f%C5%91bb%20t%C3%ADpusai%20-%20Kov%C3%A1cs%20Z.pdf
- Max Pooling* [é. n.]. Online: <https://paperswithcode.com/method/max-pooling>
- NÉMETH András – VIRÁGH Krisztián (2022): Mesterséges intelligencia és haderő – Polgári alkalmazási lehetőségek V. rész. *Haditechnika*, 56(5), 2–7. Online: <https://doi.org/10.23713/HT.56.5.01>
- NÉMETH András – VIRÁGH Krisztián (2023): Mesterséges intelligencia és haderő – Katonai alkalmazási lehetőségek VII. rész. *Haditechnika*, 57(1), 2–6. Online: <https://doi.org/10.23713/HT.57.1.01>
- Objektum felismerő (képfelismerő) AI megoldásunk* (2019). Online: www.regens.com/hu/-/objektum-felismeres-kepfelismeres-mesterseges-intelligenciaval-bemutato-video
- SHARMA, Pulkit (2019): *Image Classification vs. Object Detection vs. Image Segmentation*. Online: <https://medium.com/analytics-vidhya/image-classification-vs-object-detection-vs-image-segmentation-f36db85fe81>

Teachable Machine [é. n.]. Online: <https://teachablemachine.withgoogle.com/>

TOMOLYA János – PADÁNYI József (2012): A terrorizmus jelentette kihívások. *Hadtudomány*, 22(3–4), 34–67. Online: www.mhtt.eu/hadtudomany/2012/3_4/HT_2012_3-4_Tomolya_Padanyi.pdf

TÓTH László (2019): *Mesterséges neuronhálók és alkalmazásai*. Online: www.inf.u-szeged.hu/~tothl/ann/Neuronhalok-egyben.pdf

What Is Deep Learning? 3 Things You Need to Know [é. n.]. Online: www.mathworks.com/discovery/deep-learning.html