**NATIONAL UNIVERSITY OF PUBLIC SERVICE**

**Doctoral School of Military Engineering**

Viktor Huszár:

# Possibilities and challenges of the defensive use of artificial intelligence and computer vision-based technologies and applications in the defence sector

Doctoral (PhD) dissertation

**Supervisors**:

Dr. Imre Négyesi

. . . . . .

Dr. Csaba Krasznay

. . . . . .

**Budapest, 2023 February**

# Chapter 1

# Introduction

## 1.1 Introduction

The use of artificial intelligence (AI) and machine learning (ML) techniques in various fields has been increasing rapidly in recent years. One such field is safety and security, where these technologies are being utilized to improve the detection and prevention of potential threats. AI and ML algorithms can analyze large amounts of data in real-time and make predictions, which can be used to alert security personnel or trigger automatic responses in the case of an emergency. However, the implementation of these technologies in safety and security applications presents several scientific challenges that must be addressed for their practical deployment.

One major challenge in the use of AI and ML in safety and security is the need for accurate image and video classification. In military and security applications, accurate classification of images and videos is critical for detecting and identifying potential threats, such as weapons, explosives, and suspicious activities, in real-time. Current methods for image and video classification have limitations in terms of their reliability, especially in security applications where misclassification can have severe consequences. This challenge requires the development of more robust and accurate classification algorithms, as well as the creation of larger and more diverse datasets for training.

Another significant challenge is the need for generalizability in AI and ML models. The ability of trained models to generalize to new scenarios is crucial for their practical use. However, existing literature on safety and security applications using AI does not adequately address the generalizability of trained models, particularly through cross-database validation studies. This challenge requires the development of new evaluation techniques that can measure the generalizability of trained models across different scenarios.

Furthermore, the development of prototypes and stand-alone implementations is es-

sential for applying research findings in practical safety and security applications. Much of the current research in safety-related applications do not address the applicability of developed methods or how they can be adapted for practical use cases. Addressing this challenge will require the development of more practical and scalable solutions that can be easily integrated into existing safety and security infrastructures.

The use of AI and ML technologies in safety and security applications has the potential to significantly improve the detection and prevention of potential threats. However, the implementation of these technologies presents several scientific challenges that must be addressed for their practical deployment. Addressing these challenges will require the development of new research and innovative solutions that can enable the development of practical safety and security applications using AI.

## 1.2 Scientific challenges

Despite the potential of AI and computer vision in safety and security, their practical implementation faces several scientific challenges. These include deep learning for image/video classification, limited dataset diversity, insufficient generalizability, and lack of stand-alone prototypes. Addressing these challenges will require innovative solutions and new research to enable the development of practical safety and security applications using AI.

## 1.3 Research hypothesis

This research aims to explore the latest advancements in artificial intelligence (AI) and deep learning technologies to enhance safety measures and improve the efficiency of decision-making processes in military settings. Three key areas of research are identified and tackled, including the detection of spoof attacks, violence detection, and the ability to run highly resource-intensive AI/deep learning algorithms in a distributed environment. The hypothesis for each area is as follows:

- **H1**: Deep learning-based methods can effectively detect spoof attacks from human video frames.
- **H2**: Deep learning-based methods for action recognition can be adapted for violence detection.
- **H3**: Distributed ledger technology can effectively run high resource-intensive AI/deep learning algorithms.

These hypotheses seek to contribute to the development of more robust and effective security systems for the military.

## 1.4 Bridging research and practice in AI for defence

Digital Guard is a security system that utilizes AI, machine learning, and computer vision to detect security breaches in various settings, including military bases, airports, and critical infrastructure. The system identifies and tracks people, objects, and vehicles in real-time and generates alerts when it detects suspicious activity. This research aims to explore the latest advancements in computer vision and AI technologies for safety purposes, with a focus on military bases and educational institutions. By automating the process of identifying and responding to potential threats, the Digital Guard has the potential to greatly improve the efficiency and accuracy of safety measures. The ultimate goal is to establish a proactive and efficient tool that can identify and respond to potential threats before they escalate into dangerous situations.

## 1.5 Research Methodology

This research proposes novel methods to overcome the shortcomings of existing techniques and advance state-of-the-art results. The first two key areas of research focus on improving spoofing and violence detection through the development of reliable and effective tools using state-of-the-art deep learning algorithms. The research on distributed computing explores how distributed ledger technology can improve the scalability and efficiency of AI/deep learning algorithms. A two-stage approach was undertaken, with a literature review in the first stage and the proposal of novel methods in the second stage. The new scientific results are detailed in the subsections. The overall methodology involves a combination of empirical and applied research approaches to develop innovative and practical solutions to challenges in the field of AI for security and military applications.

# Chapter 2

# Summary of new scientific results

The new scientific results presented in this dissertation have significant implications for the defence sector, AI and computer vision (CV) researchers, and policymakers shaping the future of these technologies. While the results do not directly contribute to the field of machine learning and artificial intelligence, they provide valuable insights into the development of AI-based solutions for real-world challenges. Specifically, the novel approaches to spoof detection and violence detection using deep learning techniques, as well as the use of Distributed Ledger Technology for decentralized processing of large datasets, have the potential to significantly improve public safety and security. Below, these new scientific results are presented in a concise numbered list and the main findings and contributions from this work are detailed in the following subsections.

1. I discovered that deep learning-based methods are highly effective in detecting spoof attacks in human video frames, providing a practical tool for identifying fake video footage that fools smart digital systems.

2. I created a novel approach to violence detection using deep learning-based methods for real-time detection of violent behaviour, with the potential to significantly improve public safety and security.

3. I investigated the use of distributed ledger technology (DLT) for the decentralized processing of large datasets and high computational tasks, providing insights for improving the scalability and efficiency of AI and deep learning algorithms in military contexts.

## 2.1 Thesis Group I - Live spoof detection for Human Activity Recognition (HAR) applications.

- **I created a deep learning architecture for detecting spoofing attacks in videos using replay techniques.**
- **I discovered the effectiveness and generalizability of the proposed architecture through comprehensive evaluation and testing.**

**Relevant publications: [1] [2]**

### 2.1.1 Novel deep learning architecture for spoofing detection

*I created an ensemble multi-stream model that detects spoofing cases arising from video replay attacks. The model inspects distinct regions of human faces and combines these observations to provide robust classification between spoof and genuine cases.*

- Frames are extracted and fed to the pre-trained YOLO model to extract windows corresponding to the human head region on the image. Each detected head bounding box on a video frame is resized to 64 × 64 × 3 pixels and processed in three different streams that use CNN architecture inspired by VGG16 net as a baseline model. The output of the three streams is combined using the obtained classification probabilities following majority voting to classify spoof and genuine cases. The three streams process:
  - Predicted and resized head bounding box of size $64 \times 64$ pixels.
  - Cropped lower $64 \times 32$ pixels from the resized image.
  - Cropped central $32 \times 32$ pixels from the resized image.
- The effectiveness of the proposed ensemble method is assessed using state-of-the-art methods in the literature for spoofing detection applications using facial image data: Local Binary Pattern (LBP) histograms, Local Binary Pattern histograms from Three Orthogonal Planes (LBP-TOP), Statistical Binary Pattern histograms (SBP) and Statistical Binary Pattern histograms from Three Orthogonal Planes (SBP-TOP). Experimental results show that the proposed EM method outperformed other methods considerably.

### 2.1.2 Temporal analysis for real-time spoofing detection

*I formulated a strategy to combine multiple detections of the proposed deep learning network temporally on a captured video or on a live video stream to systematically detect video replay spoof attacks while maintaining real-time performance.*

- Chunks of overlapping video clips, each containing several frames (depending on the model under testing) with an overlap of 15 frames are defined.
- Within a considered clip, the proposed deep learning models are run on three frames that are 15 frames apart and the obtained predictions are combined to derive a single prediction for this clip. Following this method, the first prediction is made after 31 frames and thereafter, predictions are made every 15 frames.

### 2.1.3 New database for spoofing detection

*I created a new database comprising real and spoof videos captured from 38 different users in different locations and under different lighting conditions. The database is diverse and precisely captures the required features for training the proposed deep learning network.*

- A diverse database consisting of almost 50,500 full HD images of 38 users (both male and female) between the ages of 8 and 40 juggling a football in different backgrounds and lighting conditions are collected. These are extracted from several videos shot using various iPhones — iPhone 6, 6S, SE, 7, 8, X, and XS.
- These images are manually labelled with bounding boxes that encapsulate the following data: human body parts — head, left shoulder, right shoulder, left elbow, right elbow, left hand, right hand, left hip, right hip, left knee, right knee, left foot, right foot and also the bounding box of the football.
- YOLO deep learning CNN architecture is trained using the captured and labelled database and used for detecting the required body parts together with the ball from the video frames in a single shot.
- This database is also used to learn the genuine and spoof cases. To this end, an additional 50,500 Full HD spoofed images were generated using the original images by capturing the same videos on several monitors: a 27-inch Dell 4K monitor, a 15-inch Full HD Lenovo laptop monitor, and a 13- inch MacBook Pro monitor with a resolution of 2560 × 1600.
- Before training the networks, all video frames are pre-processed to extract the head region using the previously trained YOLO model and resize the head image to 64 × 64 pixels.
- Data augmentation techniques are also used to reduce problems from overfitting including increasing and decreasing the pixel brightness by a value of 50, doubling

and halving the image contrast, and adding Gaussian noise with variances of 50 and 100.

- The pre-processing steps were consistently added to all the video frames. Finally, before training, data was also shuffled. Further, all the pixel intensities are scaled to the range [0, 1] for training.

### 2.1.4 Analysis about model generalizability

*I created empirical evidence through additional evaluation of the performance of the proposed approach in the context of biometric recognition applications, demonstrating its applicability in other domains.*

- To evaluate the ability of the proposed spoofing detection model to generalize to other domains such as face recognition systems, I have experimented with two widely known datasets in this context: Idiap REPLAY-MOBILE and CASIA Face AntiSpoofing.
  - The REPLAY-MOBILE dataset consists of 1190 video clips of photo and video presentation attacks (spoofing attacks) to 40 clients, under different lighting conditions. These videos were recorded with an iPad Mini2 (running iOS) and an LG-G4 smartphone (running Android) in full HD resolution.
  - The CASIA Face AntiSpoofing Database consists of 600 video clips of 50 subjects. Out of the 600 video clips, 150 clips represent video replay attacks. Compared to the Idiap database, the CASIA DB provides images captured using a variety of cameras (Sony NEX-5-HD, two low-quality USBs) to capture replay attacks displayed on an iPad. However, a significant deficiency of this database is that the video replay attacks are captured in very low resolution ($640 \times 480$).
- Results show that the proposed ensembled approach performs very well on the REPLAY-MOBILE database irrespective of the fact that the users are located very close to the camera (higher IPD values than that of my database).
- Even though, the CASIA database consists of low-resolution images, the proposed method performed reasonably well also on this database.

### 2.1.5 Prototyping and stand-alone implementation

*I created a solution by designing an IOS mobile application that implements the proposed approach in real-time, bringing the technology to the convenience of mobile devices.*

- I have implemented the proposed Ensemble multi-stream model in swift for IOS. The application is standalone and tested on iPhone 8 released in 2017 and has a 2.39 GHz hexacore 64-bit.
- The developed models as well as the YOLO model for head bounding box detection are converted to CoreML API for running on the IOS device. The application takes as input incoming frames from the on-device-camera and runs my trained YOLO model to detect head bounding box information.
- The algorithm works in real-time. The proposed method can run on an iPhone 8 and processes a single frame on an average of 4 ms. This translates to a frame rate of about 250 frames per second.
- Following the proposed testing scheme, running the proposed model every 15 frames further ensures that there are enough resources left on the device to run the activity recognition applications.

### 2.1.6 Experiments with video compression

*I created empirical evidence by simulating and evaluating the performance of the proposed ensemble model under extreme video compression (300 kbps) and discovered its robustness.*

- Video compression techniques are applied to reduce the video bitrate to facilitate efficient streaming which introduces video artefacts. To evaluate the ability of the proposed ensemble model to correctly classify the spoof cases, I have generated compressed video streams with varying bitrates from 300 kbps to 1500 kbps. Multiple videos were generated with different bitrates using FFmpeg to experiment with video compression.

## 2.2 Thesis Group II - Violence Detection for automated video surveillance applications

- **I created a solution for automatic violence detection in video surveillance by exploring smart networks.**
- **I discovered a novel deep learning-based approach for violence detection that outperforms existing methods with fewer model parameters and demonstrated robust performance under compression artefacts commonly encountered in remote server processing.**

**Relevant publications: [3]**

### 2.2.1 Efficient deep learning architecture for violence detection

*I created a deep learning-based method that can be used to filter violent and normal patterns videos. Considering the popular video classification metrics for evaluation, the method outperforms several state-of-the-art methods for violence detection and is also able to cope with video compression artefacts, while remaining computationally lightweight.*

- I have explored X3D-M deep learning architecture that is computationally lightweight to learn and detect violence patterns from videos. I proposed two architectures - fine-tuned and transfer-learned models for classifying video clips containing violence, which leverage action recognition features learned from the Kinetics-400 dataset. For both models, I have modified the architecture into a regression model to generate a violence coefficient that indicates the probability of the existence of violence in a given video clip.
  - The fine-tuned model trains all the parameters of the adapted X3D-M model on the datasets for violence detection. The second fully connected layer of the X3D-M model is replaced to output a floating point variable which is converted into range $[0, 1]$ using a sigmoid function to derive the violence coefficient.
  - The transfer learned model uses X3D-M for inferring video features and does not retrain the parameters of the original X3D-M model. Pre-processed videos containing both violence and no violence are inputted into a pre-trained X3D-M model for feature extraction. Three additional fully connected layers are trained using the extracted features to obtain the violence coefficient.
- The experimental results on individual datasets show that the fine-tuned model performed better than the state-of-the-art methods on most datasets with relatively fewer model parameters.
- Transfer learned model also achieved decent performance on all the datasets given that, it has less trainable parameters than fine-tuned model and thus relatively less adaptable to specific scenarios.
- When testing on all combined datasets, the fine-tuned model achieved better performance and the transfer learned model produced more combined false positives and false negatives.

- Further tests on individual datasets show that models trained on all combined datasets did not perform well in several cases when compared to the performance of models trained on individual datasets. This shows that there are multiple inconsistencies in the publicly available datasets for violence detection.

### 2.2.2 Comprehensive database for violence detection

*I created a comprehensive database of violent and normal videos by combining and extending seven existing video databases, providing a robust resource for violence detection research across various contexts.*

- For experimenting with violence detection and to facilitate comparing the results with other methods, I have considered seven different datasets that are commonly used in literature. I have also extended some of the datasets with annotations to assist in-depth cross-validation experiments. These are described in the following:
  - **Crowd Violence** (CV) dataset contains videos involving violence in crowds, collected from YouTube.
  - **Hockey Fights** (HF) dataset is a collection of fights between players in hockey games from the USA's National Hockey League (NHL).
  - **Movie Fights** (MF) dataset collects several scenes from action movies.
  - **Real Life Violence Situations** (RLVS) dataset gathered fighting videos from YouTube and also from real street cameras that contain many real street fights.
  - **Real-World Fight-2000** (RWF-2K) dataset is a collection of large-scale fighting videos from YouTube. The dataset contains trimmed video clips captured by surveillance cameras from real-world scenes.
  - **UCF-Crime Selected** (UCFS) dataset is a subset of the UCF-Crime dataset. The UCF-Crime dataset contains long untrimmed surveillance videos that cover 13 real-world anomalies including Abuse, Arrest, Arson, Assault, Burglary, Explosion, Fighting, Road Accident, Robbery, Shooting, Stealing, Shoplifting, and Vandalism without annotations. Although this is a large-scale dataset, all videos in the violence class contain a mix of violent and normal actions which is undesirable. Among the anomalies, I selected the classes - Abuse, Explosion, Fighting, Road Accident, and Shooting and manually trimmed these videos to only contain violent parts for training and testing.
  - **XD-Violence Selected** (XD-V) dataset contains a subset of videos from the XD-Violence dataset. XD-Violence dataset contains several untrimmed videos covering 6 anomalies including Abuse, Car Accidents, Explosions, Fighting,

Riots, and Shooting gathered from action movies and YouTube. Similar to the UCF-Crime dataset, I selected a set of videos belonging to the classes - Abuse, Explosion, Fighting, Road Accident, and Shooting and manually trimmed these videos to only contain violent parts for training and testing.

- All datasets also contain normal videos for training and testing that have no violence involved. In the case of UCFS and XD-V datasets, normal videos are trimmed to five-second video clips to match the average duration of normal clips of other datasets. Also in the case of UCFS and XD-V, the maximum duration of a video clip containing violence is limited to approximately 5 seconds.

### 2.2.3 Analysis on the generalizability of proposed models

*I discovered high-performance results through a thorough evaluation of the proposed approach on a comprehensive database of violent and normal videos, including cross-database validation to assess the generalizability of the methods.*

- According to the metric scores trained fine-tuned and transfer-learned models on the CV dataset did not generalize well to other datasets. This is anticipated since CV contains only examples of mass violence and the other datasets do not contain plenty of such examples. Also, the trained FT model on the HF dataset has poorly generalized to other datasets indicating that the HF dataset does not contain diverse examples of violence and contains monotonous fighting videos between hockey players. However, the TL model trained on this dataset showed generalized better than the FT model as indicated by the metric scores.

- Both models trained individually on datasets - MF, RLVS, RWF-2K, UCFS & XD-V performed satisfactorily in the cross-validation tests and generalized decently to other datasets with average accuracy scores close to or above 80% and average AUC scores close to or above 0.8. Considering both metrics, models trained on UCFS and XD-V datasets exhibited the best generalization ability in the cross-validation studies. This indicates that the datasets, that are gathered in this study, have the most representative and heterogeneous samples for actions involving violence and non-violence.

- Overall, the transfer learned model showed a better capability to generalize and has less standard deviation within testing accuracy scores for individual datasets when compared to the fine-tuned model.

### 2.2.4 System implementation

*I discovered the high performance of the developed approach through extensive evaluation on collected databases, including cross-database validation to assess the generalizability of the methods. I created a standalone functional system for automated violence detection, implementing the proposed methods using the PyTorch deep learning library. The resulting application can easily be adapted for use in surveillance applications.*

- From the incoming video stream, non-overlapping video segments having a duration of four seconds are extracted. From each segment 16 video frames are extracted into a block following uniform temporal sampling. These blocks are pre-processed and then used as input to the trained models to get a violence coefficient for the current segment.
- This application is implemented on the Ubuntu Linux operating system using AMD Ryzen Threadripper 1950X 16-core processor. I have used Nvidia GeForce GTX 1080 Ti GPU using CUDA toolbox for running my trained PyTorch models.
- Implementation results show that together with block extraction and pre-processing, for each four-second video segment, both models require $0.06$ seconds on average to infer a violence coefficient. The pre-processing is implemented on the CPU and this consumes $0.04$ seconds on average. Therefore, the average time of running the proposed model is $0.02$ seconds.

### 2.2.5 Video compression experiments

*I uncovered the impact of video compression artefacts commonly encountered in video streaming fields on the performance of proposed models through comprehensive experiments and presented the results.*

- For experiments, I have generated compressed video streams with varying bit-rates - 300, 500, 1000 & 1500 Kbps. Two datasets - RWF-2K and CV are used for the experiment and corresponding testing videos from these two datasets are compressed. Multiple videos are generated with the considered bit rates using FFmpeg.
- Results from the study pointed out that the proposed models did not show greater fluctuations in the performance and performed decently even under extreme compression (300 Kbps). This shows that the proposed models did not model the noise in the training videos and were precisely directed to learn the concept of violence.

## 2.3 Thesis Group III - Applicability of DLT and blockchain technologies in military

**I created a system that leverages DLT technologies with a focus on blockchain as a computational data resource for training machine learning and deep learning algorithms while prioritizing data privacy through the use of specialized neural networks. Relevant publications: [4] [5] [6] [7] [8]**

### 2.3.1 Data privacy, sharing, and tracking

*I designed a concept where blockchain can be used to secure and encrypt data used for deep learning, making it more difficult for unauthorized parties to access or tamper with the data.*

- Data privacy is ensured by using blockchain to encrypt and secure sensitive data, such as intelligence or surveillance data, that is collected by military assets. The data can be stored on the blockchain and is only accessible to authorized parties with the appropriate encryption keys. This can help to prevent unauthorized access or breaches of sensitive data.
- Secured data sharing in real-time between multiple parties, such as different military units or coalition partners, can also be facilitated by the use of blockchain to improve situational awareness and decision-making capabilities.
- The blockchain can also be used to track the origin and handling of data, providing transparency and accountability. For example, the blockchain can be used to record the provenance of intelligence data, such as the sources, handling, and analysis. This allows for the verification of data authenticity and integrity, and can also aid in the investigation of data breaches.

### 2.3.2 Model training and decentralized decision making

*I formulated a concept where blockchain can be used to train machine learning models in a decentralized way, using distributed computing power and data, providing a more secure and transparent way of training large and complex models. While running the trained deep learning models, blockchain-based smart contracts can enable decentralized decision-making based on consensus among multiple parties.*

- Blockchain can be used to secure and track the data used to train machine learning models for military applications. The data collected by military sensors can be

encrypted and stored on the blockchain, ensuring that only authorized parties have access to it. Additionally, distributed computing power and data can be used to train models in a decentralized way, providing a more secure and private way of training models.

- Decentralized decision-making using smart contracts can be used to ensure that actions taken by military assets, such as unmanned aerial vehicles (UAVs) or autonomous weapons systems, are based on a consensus among multiple parties. For example, a smart contract could be used to ensure that a UAV only fires a weapon if a consensus is reached among multiple operators, or if certain predetermined conditions are met. This can help to prevent unauthorized or accidental use of weapons, and can also improve the overall situational awareness and decision-making capabilities of military assets.

# Chapter 3

# Applications of the Work

The research on safety and security applications of AI and computer vision has numerous practical applications across a wide range of domains. One of the primary applications of this research is in the field of military and defence. Military installations, such as bases and airports, have unique safety challenges that require constant vigilance and proactive measures to prevent potentially dangerous situations from arising. The use of AI for safety applications in these settings has the potential to greatly improve the efficiency and accuracy of safety measures by automating the process of identifying and responding to potential threats.

Another application of this research is in the field of transportation. Airports, seaports, and train stations are all high-traffic areas that require constant monitoring to ensure the safety and security of travellers. AI and computer vision technologies can be used to identify and track people, objects, and vehicles in real-time, detecting suspicious activity or behaviour and generating alerts and notifications to security personnel. Beyond military and transportation applications, this research can also be applied in the context of public safety. For example, AI and computer vision can be used to monitor crowded areas such as sports stadiums, concert venues, and public parks to detect and prevent potential threats. These applications can help law enforcement agencies to more effectively monitor large crowds and improve public safety.

This research also applies to industrial safety and education. AI and computer vision can monitor high-risk environments, such as power plants and manufacturing facilities, to detect safety hazards in real-time. In educational institutions, they can detect safety threats like unauthorized access, improving overall safety and security.

# Chapter 4

# Conclusions and Recommendations

This research has explored the potential of using artificial intelligence and computer vision technologies for safety and security applications. Through a review of existing literature and a series of experiments, the work evaluates and provides novel solutions to the several scientific challenges identified in this field and the work also addresses the practical implementation of these technologies in real-world scenarios. The specific challenges include improving the accuracy and reliability of image and video classification in security applications, addressing the lack of diversity in training datasets, improving the generalizability of trained AI models, and developing prototypes and stand-alone implementations for practical use cases.

The outcomes of this work recommend future research efforts to focus on developing innovative solutions that involve the creation of more diverse and representative training datasets, the integration of transfer learning and other techniques to improve model generalizability, and the development of practical prototypes and implementations that can be tested in real-world scenarios. In addition, the research also recommends that future research in this area should focus on developing ethical guidelines and regulatory frameworks that ensure the responsible use of these technologies in safety and security applications. This could include guidelines for data collection, use, and storage, as well as guidelines for the development and deployment of AI models in safety-critical applications.

Overall, the findings of this research suggest that artificial intelligence and computer vision technologies have significant potential for improving safety and security in a range of applications, from military bases and airports to critical infrastructure and educational institutions. However, to fully realize this potential, we must overcome the identified challenges that currently limit the practical implementation of these technologies and develop ethical and regulatory frameworks that ensure their responsible use.

# List of Publications


[1] Viktor Dénes Huszár and Vamsi Kiran Adhikarla. Live spoofing detection for automatic human activity recognition applications. *Sensors*, 21(21):7339, 2021.

[2] Viktor Huszár. Hamisítás észlelési módszerek az automatizált emberi tevékenység felismerésben. *Hadtudomány*, 32(1):270–284, 2022.

[3] Viktor Dènes Huszár, Vamsi Kiran Adhikarla, Imre Négyesi, and Csaba Krasznay. Toward fast and accurate violence detection for automated video surveillance applications. *IEEE Access*, 11:18772–18793, 2023.

[4] Huszár Viktor. A blokklánc, a számítógépes látás és a mesterséges intelligencia alkalmazási lehetőségei a kiberhadviselésben. *Hausner, Gábor (szerk.) Szemelvények a katonai műszaki tudományok eredményeiből II., Ludovika Egyetemi Kiadó*, 2021.

[5] Viktor Huszár. Kiberbiztonság mint a haderőfejlesztés kiemelt területe: a decentralizáció és a blokklánc-technológia lehetőségei a kibertérben. *Honvédségi Szemle: A Magyar Honvédség Központi Folyóirata*, 148(3):3–17, 2020.

[6] Huszár Viktor. A decentralizáció és a blockchain-technológia felhasználási lehetőségei gépi látás és mesterséges intelligencia használatával a katonai szervezetekben. *Hadmérnök*, 14(4):179–189, 2020.

[7] Viktor Huszár. Distributed intelligence: Cyber warfare application possibilities of computer vision and artificial intelligence. *Honvédségi Szemle–Hungarian Defence Review*, 149(1-2.):4–19, 2021.

[8] V. Huszár. Application possibilities of decentralization and blockchain technology using computer vision and artificial intelligence in defense management, military and police organizations. *Honvédségi Szemle: A Magyar Honvédség Központi Folyóirata*, 148(1.-SI):4–14, 2020.

# CURRICULUM VITAE

Address: Vérhalom utca 17-19. E
LPH. 3. em. 1 a.
1025 Budapest Hungary
Tel.: +36(06)303106944

Email: viktor.denes.huszar@uni-nke.hu
Website: https://viktorhuszar.com

Date of birth: 8 May 1985
Place of birth: Budapest, Hungary
Nationality: Hungarian
Civil status: Married, three children

## RESEARCH INTERESTS

Artificial Intelligence, Computer Vision, Blockchain, Machine Learning

## EDUCATION

| | |
|---|---|
| 2019 - present | Active Researcher, National University of Public Service, Doctoral School of Military Sciences |
| 2020 – 2021 | Certificate in Company Direction, Institute of directors, (IoD), UK |
| 2014 – 2018 | Master of Science in Security and Defence Policy, National University of Public Sciences |
| 2011 – 2012 | Producer Diploma, European Audiovisual Entrepreneurs |
| 2006 – 2009 | BA in Business Administration University of the West of England, Bristol Business School |
| 2004 – 2008 | BA in Economics, Inholland University, School of Economics |
| 2005 – 2009 | BA in Economics, Berlin School of Economics |
| 2000 - 2004 | Bilingual Diploma,International Baccalaureate Organization |

## Scholarships

- (Kooperatív Doktori Program Doktori Hallgatói Ösztöndíj (KDP) 2021/2022-es tanév)
- New National Excellenece Scholarships Programme, Semester 2021/2022
- (Új Nemzeti Kiválóság Program Ösztöndíj (ÚNKP) 2121/2022-es tanév)
- *Ministry of Interior Scholarships Programme, Semester 2021/2022*
- (Belügyminisztérium Gyakornoki Program 2021/2022-es tanév)

## Current Positions

- Chairman of FITEQ
- Co-Founder and Co-Inventor of Teqball
- President of the Football Club of the Budapest University of Technology& Economics
- Hungarian Economic Association Member
- Board member of the Foundation for Children Health Education and Prevention
- World Aquatics Reform Committee and Digital Sub-committee member

## PUBLICATIONS

- Hamisítás észlelési módszerek az automatizált emberi tevékenység felismerésben, Nemzeti Közszolgálati Egyetem, Katonai Műszaki Doktori Iskola, Évf. 32 szám 1 2022: Hadtudomány.
- A blokklánc, a számítógépes látás és mesterséges intelligencia alkalmazási lehetőségei a kiberhadviselésben. In: Hausner Gábor (szerk.) Szemelvények a katonai műszaki tudományok eredményeiből II., Ludovika Egyetemi Kiadó *2021, pp. 147-161.*
- Live Spoofing Detection for Automatic Human Activity Recognition Applications, *SENSORS 2021 pp. 1-20.*
- Methods and systems for performing object detection and object/user interaction to assess user performance, *Szabadalom (IP-20210902)*
- Kiberbiztonság mint a haderőfejlesztés kiemelt területe: a decentralizáció és a blokklánc technológia lehetőségei a kibertérben, *Honvédségi Szemle 2020/3., p3-17.*
- Application Possibilities of Decentralization and Blockchain Technology Using Computer Vision and Artificial Intelligence in Defense Management, Military and Police Organizations, Hungarian Defence Review,*Special Issue 2020, Vol. 148, Nr. 1 DOI* https://doi.org/10.35926/HDR.2020.1.1
- Javított tulajdonságú többcélú sporteszköz, *Oltalmi formák/ Magyar szabadalom, 2020*
- Application possibilities of decentralization and blockchain technology, American Journal of Research, Education and Development, *Special Issue 2020, Vol. 148, Nr. 1*
- Cyber Warfare Application Possibilities of Computer Vision and Artificial Intelligence, ICCECIP 2020 Absract book Budapest, Óbudai Egyetem, ISBN 978-963-449-221-4
- A decentralizáció és a blockchain technológia felhasználási lehetőségei gépi látás és mesterséges intelligencia használatával a katonai szervezetekben, *Hadmérnök 2019, 14. évfolyam, p179-189*

## Conference Presentations

- Applications of Artificial Intelligence in Human Activity Recognition (ICCECIP 2021), NKE Ludovika, 15.11.2022
- Live Spoofing Detection for Automatic Human Activity Recognition Applications (Military Science and Military Art International Thematic Conference), Ludovika – University of Public Service, Hungary, 21.10.2022.
- Distruptive technoligies (ICT Sping 2021), Luxemburg, 17.09.2021

## Awards

- Recognition of Teqball Sport as Hungarikum (2022), Industrial Innovation Award Teq LITE Table (2020),
- Angéla Németh Medal for Hungarian, University Sports Contributions (2020),
- Red Dot Design Award Teq LITE table (2020),
- Design Management Adward (2019),
- Hungarian Design Award Teq Smart Table- (2019),
- iF Design Award Teq SMART table (2018),
- ISPO Award Teq ONE table 2015/2016,
- Red Dot Design Award Teq ONE table (2015),
- Hungarian Futsal Championship Gold medal (2013)

## SKILLS

- Languages: Hungarian (native), English (fluent), German (advanced), Dutch (intermediate) , Spanish(intermediate) , NATO STANAG 6001 (Military English, Level 3)
- Software: LaTeX, MS Office, Linux/Debian