Modular augmented reality platform for smart operator in production environment

Jumyung Um
Technologie-Initative SmartFactoryKL
Kaiserslautern, Germany
jumyung.um@smartfactory.de

Jens Popper
Technologie-Initative SmartFactoryKL
Kaiserslautern, Germany
popper@smartfactory.de

Martin Ruskowski
IFS - Innovative Fabriksysteme
Deutsches Forschungszentrum für
Künstliche Intelligenz
Kaiserslautern, Germany
Martin.Ruskowski@dfki.de

Abstract— **The trend of mass customization requires highly flexible production systems, which pose a new challenge for workers. Numerous concepts using augmented reality technologies have been developed to offer support in manual assembly processes. Newer AR devices are equipped with more and more functions, such as microphones, speakers or multiple camera systems, which can be used to help workers. However, wearable devices offer only limited computational power and limited flexibility of their interface platform. This poses a challenge to heavy data transactions from various different sources and restrict the use for real-time responses, which are often necessary for applications in a factory. To overcome this limitation, the author of this paper proposes a combination of wearable devices with edge devices located on the shop floor. An augmented reality platform was constructed with modularized services by using the digital twin of the machine components and machine learning algorithms for object recognition. This is integrated in the factory environment to synchronize real machine components with their digital twin.**

Keywords—Augmented reality, Smart Factory, Computer vision, Machine Learning, Digital Twin, Edge computing

I. INTRODUCTION

Change in market demand leads to more challenges for factory workers. They need to be able to handle more products and complex machinery than ever before. To handle that, augmented technologies can help them to gain the required proficiency. Many concepts in existing literature propose an assistance system which uses tablet pcs or projectors. These, however restrict the worker in their freedom of movement or require them to hold additional devices in their hands while working. A way to overcome these shortcomings is to utilize wearable devices, such as smart glasses.

The use of cloud platforms for complex calculations, for example, in the field of machine learning, is not a viable solution either. Company regulations often prohibit the use of raw data outside the factory. Furthermore, the communication with a cloud takes some time, which limits its application for time-sensitive information.

Even if cloud services are available, the operation platforms of commercial smart glasses follow vender-specific interface and protocol because there is not standardized platform in AR industry now. It leads to the problem of interoperability and flexibility of various user interfaces such as dashboard monitor, smart watch/phone/glasses. User interfaces and communication protocols are not interchangeable.

Smart glasses pose some technical challenges when used in a production line. Due to their small size and the requirement to be light-weight, they often possess limited computational power and a small battery capacity. Furthermore, they often do not have large data storage. Therefore, most commercially available smart glasses are not suitable for computationally intensive applications.

This paper proposes a way to overcome these limitations. It uses edge computing to compensate the disadvantages of limited computing power, data storage and modularization. The edge node updates the digital twin of the equipment in real time and provides the necessary guide videos to the operator. At the same time, the edge node provides an algorithm platform for machine learning and computer vision.

Based on the proposed concept, the paper explains the architecture to give the augmented performances to smart glasses by operating the machine learning algorithm and the video storage communicating with human user interface. Based on this, the proposed concept is implemented in the SmartFactory for testing purposes and different options for wireless data transmission are evaluated.

II. LITERATURE SURVEY

The application of wearable AR devices in various application fields has been surveyed in [1]. Applications which enable augmented reality applications in manufacturing processes have been evaluated in [2]. It proposed to consider "Reliability", "Responsiveness" and "Agility" as central criterion to evaluate the applicability of AR devices. Other important criteria are "User-friendliness" and "Scalability" [3].

Many application fields for AR in the manufacturing context have been identified, such as maintenance [4], assembly [5] or logistics [6]. One of the most common fields for AR in manufacturing is assistance systems. The first application of AR for worker assistance has been presented in [7]. It assisted in a manual assembly process by giving step-by-step instructions. Many assistance systems for workers utilize AR on a workstation, such as [8]. Furthermore, the support of workers in rapidly changing production environments has been evaluated in [9]. It showed the promising capabilities of AR to be used as a guidance system in a factory.

However, most of the presented concepts for worker support via AR heavily relies on data and processing in the AR device itself or are fixed to a set workstation.

Utilizing the edge layer for video streaming has been proposed in several papers [10, 11].

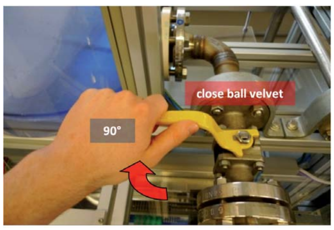
The method proposed in this paper is based on an augmented reality device that communicates with a server at edge level. This combination is used to provide the worker with multimedia supported guidance.

III. Methodology

A. *Universal Interaction devices in the SmartFactory*

The architecture platform to support human operators has been developed by SmartFactoryKL through various research and industrial projects, and have been defined as the concept of "Universal interaction device" as shown in Figure 1. This concept deals with numerous wireless technologies such as Bluetooth, WLAN, RFID, and NFC, and has the capability of connecting various wired networks and devices by cable, employing, for example, RS232, Profibus, and USB. The first prototype UCP450 merges all these communication technologies' potentials in a user-oriented way and thereby allow the users to remotely access devices, plant modules, products, product components, and order management in an easy and convenient way [12].

a) Universe interaction device, UCP450, to control whole modular factory



b) is smart glasses application to guide maintenance works

Figure 1. The applications of universal interaction device of SmartfactoryKL

After Industry 4.0 influences the factory network system, Cyber-Physical System let manufacturing resources connect with others. Since many manufacturing resources have been integrated with each other, new requirement of universal interaction devices is to cope with huge amount of communications with controller, sensors, user device as well as video streaming servers to providing visual guideline to human operators through wireless connections. For this reason, new platform architecture is required.

B. *Platform architecture*

In the production line, there are plenty of things that need to be controlled and help to boost up the speed of Human interaction. In the AUTOWARE project funded by EU H2020 project, the smart glasses are used to adding the mixed reality in the industrial environment. Object detection base on our specific training data not only is helpful to monitor what is important for safety service, but also represent guide related to the specific situation and task like manual assembly.

The proposed platform architecture of this paper consists of a set of smart glasses and the edge server connected with production module as shown in Figure 2.

Applications, which run locally on the smart glasses, are using the cameras and distance sensors for gesture recognition and object tracking. Data from these sensors is also transmitted wirelessly to the edge layer, which provides a platform for instruction videos and computationally intensive applications. This way, the images captured by the cameras of the smart glasses can be streamed to the edge server, where it is processed. Instructions can then be streamed back to the display of the glasses.

Local applications on the edge server are used for object recognition. Furthermore, the visualization of the digital twin is synchronized with the streamed video information. These applications are implemented on an edge server, since they require a high level of processing power and energy.

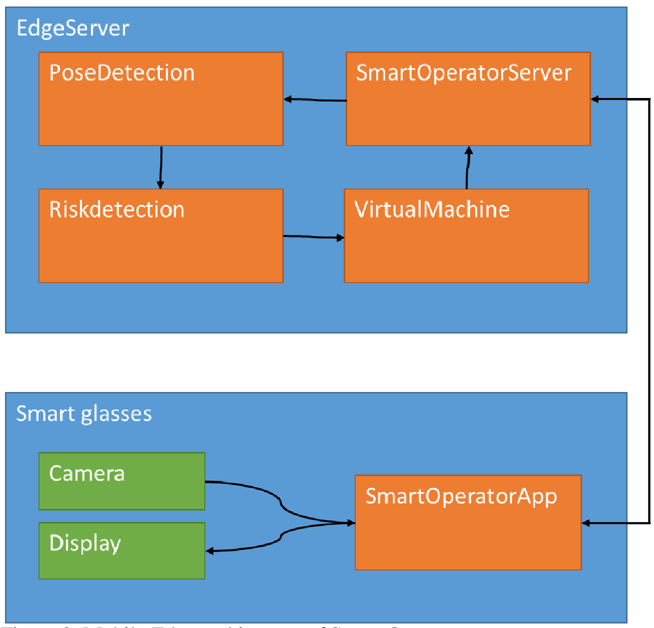


Figure 2. Mobile-Edge architecture of Smart Operator system

Main functionalities are the security of human operators in factory environment with pose detection algorithm and risk detection algorithm as well as the guide system helping in product-manipulation (e.g. assembling). Main inputs camera image of smart glasses and image of camera captured in production line. Using Object detection in python (base on Tensorflow) to detect the objects base on our specific data set. Transferring Image from Hololens and cameras in the environment to the Object detection and get the result base on which object is detected and specific tasks (Persons crossing red lines, Risky situation, Components of product, Assembly particles, etc.). The output would vary based on the functionality. That would be warning in the unsecure situation for security reason. Represent detected components of the product and help to see all particles in detail and give the overview at any part of the task process.

C. *Data transfer between the smart glasses and the edge server*

Design considerations include the speed of updates and the scalability with peripheral information. In order to react to the operator's movement in real time, fast upload speeds are needed. It should be possible to accommodate the changes in the factory environment as much as possible and show them to smart glass.

Smart glasses will conduct the update action of predefined virtual model of information received. Currently, the Smart glasses app includes 3D model of production modules to reduce time of synchronizing digital twin. Further approach to increase the flexibility is to use dynamic 3D module system. And the model updated by visualization server providing the 3d model along the target system which human operator observing now.

The edge server stores the virtual model of a factory environment and continuously update its status to receive the real-time message from production modules and sensors. User input obtained by smart glasses is transferred into edge server through wireless communication. The input activates the response message of updated virtual model. The interface function takes some time for response time.

D. *Data Processing in the Edge Server*

The design considerations of the data processing in edge serve are short running time of analysis algorithm, the integration with other sensor data, and the hand-over of high computing work into edge server. Computer vision and sensor fusion require complex algorithms to achieve valuable analysis. Therefore, the smart glass app reduces the burden of algorithm computation by embedding the heavy algorithms, such as image processing and machine learning, into the edge server. As shown in Figure 3.

The worker's camera transmits continuous streaming data to the edge server via the smart glass app. This data is analyzed by computer vision algorithm in order to detect the presence and location of objects and humans. This algorithm processes the vast amounts of data to operate image processing and neural network. Therefore, high computing power is required.

In addition, integrated data processing is possible, and each production module has own camera attached inside, and the object recognition algorithm is connected with the camera too. In the proposed architecture, the presence of human in front of a production module is displayed in the smart glasses in the place where smart glasses camera cannot see through. The integrated information processing transmits the all changes into digital twin synchronizing with human operator's user interfaces such as smart glasses.

Data processing of edge server side are composed of streaming thread, image processing thread and neural network thread. Captured image is arrived as streaming message, converted into image frame,

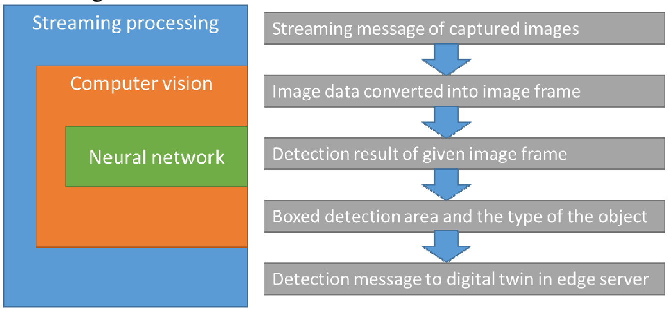


Figure 3. Data processing of Object detection in Edge server

IV. Case Study

A. *Implementation Environment*

The smart factory consists of the 11 production modules, which is shown in Figure 4, that allows all kinds of layout sequences and are detachable as needed to enable high flexibility. In these circumstances, it takes a lot of time for workers to stabilize production line. Therefore, mobile user device, such as tablet, smart watch and smart glasses, are utilized to provide the guide system of human operators. Equipping intelligent services, mobile user device need the support of high computing power with low battery consumption.

Figure 4. Modularized production system where the proposed platform applied

In the production line there are plenty of things that need to be controlled and help to boost up the speed of Human interaction as show in Figure 5 a). Object detection base on the specific training data not only is helpful to monitor what is important for safety service, but also represent guide related to the specific situation and task like manual assembly with digital twin depicted in Figure 5 b).

The deployed system was used as a smart glass tool for Microsoft. It is equipped with a camera and a distance sensor, which drives an app that is deployed with a. NET deployment. The edge server uses wireless LAN to communicate with smart glasses. Machine learning algorithm developed in

Tensorflow while image processing algorithm in OpenCV. Official training data set of object detection is used for the training of neural network graph written Tensorflow. The specification of experiment device is described in Figure 6.

a) Human operator handling digital twin of production module

b) Digital twin with physical objects in Smart glasses

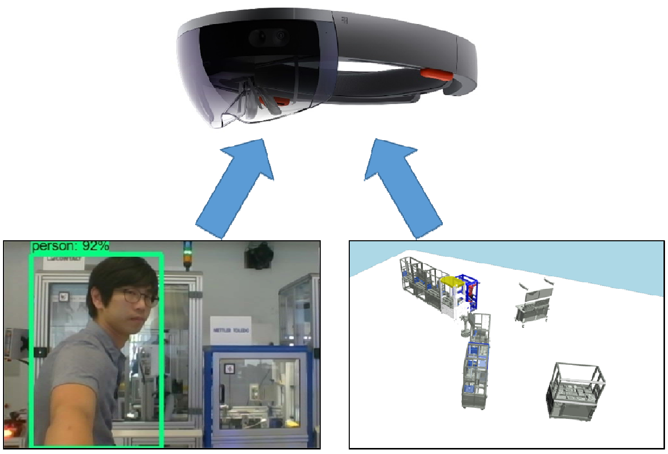
Figure 5. Smart Operator prototype displaying digital twin connected with Edge computer

Specification of Smart glasses and its emulator

MS Hololens: Intel 32 bits TPM 2.0, HPU 1.0, 2GB RAM

Emulator#1: Ubuntu 64, Quad core with 2 GB RAM

Emulator#1: Windows 10 64, Quad core with 4 GB RAM



Specification of Edge computing

Windows 7: Intel I7, 16GB RAM, GPU GTX920

Figure 6. Prototype specification of Smart Operator platform

This whole structure is classified into image processing, object detection model and label dataset. Smart Glasses sends the camera stream through internal webserver with 720x480 pixel resolution. OpenCV captures the stream, samples it as a frame data and converts an image from blue-green-red color space to red-green-blue color space. Pre-defined numerical arrays stores the modified image so that TensorFlow handles given camera data. This numerical array is the input of the object detection model which is Single-Shot-Multibox Detector (SSD) with MobileNet developed by Huang et al [13] in the proposed method. This detector achieves state-of-the-art performance measured on the COCO detector task as well as the deployment on mobile device with the real-time speed. The convolution neural network of SDD initialized with pre-trained label dataset. Object detection model uses MS_coco_label_map as the label dataset.

B. Results

Around 5 second is basic speed of mobile embedded system. But it tests in the virtual machine of high power computer in order to manipulate the performance of front-end hardware representing as user device such as smart glasses or hand-held device. Minimizing the communication time by internal virtual machine inside edge computing show the significant reductions up to 0.82 second to detect the object. Wired connection can secure 1 Gbps data transfer to make the test performance close to maximum case by showing 0.87 second. The interpolation into various wireless communication protocols is that 150Mbps, 1200Mbps and 4.6 Gbps are available in in IEEE 802.11n, 802.11ac, and 802.11ad respectively.

Otherwise the throughput of wireless communication is also observed for showing the trade-off of sharing computing load with the edge server. This section analyzes the result with three aspects of overall speed, computing time and communication. Total lead time of the object detection application includes camera capture, data transfer, image processing and object detection. Computing measurement is composed of process time, resource consumption and power supply. Resource consumption is the usages of CPU, RAM and data storage such as hard drive and flash memory. In terms of power supply, smart glasses use the battery while edge server connected with power cable. As table 1 show, smart glasses are faster than the usage of edge server via wireless communication even though edge server itself is six times faster. But battery consumption of smart glasses is very fast. Major delay happens in data exchange with edge server via wireless communication while process time in smart glasses and edge server is relatively small. The improvement of capture and transfer method will significant reduction of total lead time rather than advance detection algorithm. Wireless communication is major delay and variation of total lead time. IEEE 802.11ac is about 4 second faster than IEEE 802.11ngb. But IEEE 802.11ac has bigger fluctuation of communication speed than IEEE 802.11ngb. 2.4GH microwave has a weakness on the penetration of the obstacle. Edge server should be located near the production module, which there is less obstacle of the microwave of the wireless connection.

Figure 7 shows total lead time of different setting of (1) edge server alone, (2) smart glasses alone, (3) connection via 802.11ngb and (4) connection via 802.11ac. Figure 8 shows the breakdown of 802.11ngb connection which is longest lead time among 4 configurations.

In given condition of test configuration, there are the possibility to optimize the configuration by changing 1) the screen size of observing smart glasses camera in edge computing, 2) image size published by smart glasses. 3) Reducing the obstacle between the human operator and edge server.

In the systematic aspect, the modularity of each component improves functional and resource distributions. Camera capture and object detection are separated into two parts. By allocating each function into optimized device, overall performance improves. Resource distribution happens from smart glasses to edge server which is richer computing power, memory and power supply. The prototype uses raw video stream which is covert into any image processor so that the modularization increase between any smart glasses and any edge server.

Table 1. Performance comparison of various configurations

Test configuration	Smart glasses	Comm. Method	Edge computing
Smart glasses alone without edge service (virtual machine with 2 GB RAM and Quad core)	U:4.5 to 5.5s C: 43 - 99 % M: < 488 KB (44 %)	Not used	Not used
Edge server alone (Intel I7 64bit 16GB RAM NVIDIA GTX 950)	Not used	Not used	U: 0.82 to 0.87s C: 18 - 24 % M: < 600 K G: < 10%, 226-230MB
Smart glasses communicating with Edge server via 802.11ngb WIFI	Only transferring camera images	802.11n 0.4Gbps	U:: 11 to 12s C: 27 - 50 % M:550 to 650KB G: 2 to 14%, 227 to 249 MB
Smart glasses communicating with Edge server via 802.11ac WIFI	Only transferring camera images	802.11ac 0.87Gbps	U: 7 to 9s C: 28 - 51 % M:550 to 650KB G: 9 to 18%, 240 to 270 MB

Notation
U: Updating time (AVG)
C: AVG. of CPU
M: Memory
G: GPU load, Graphic Memory

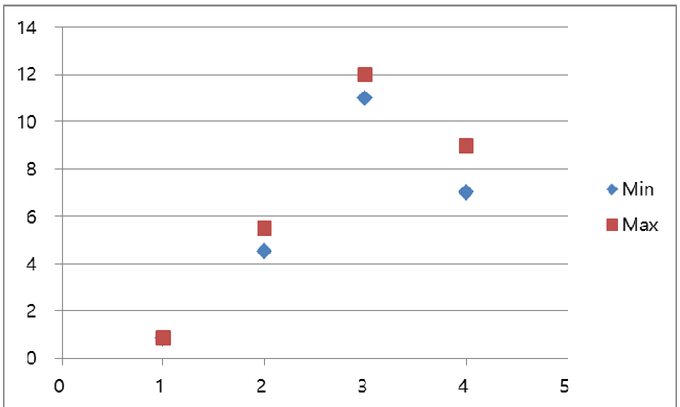


Figure 7. Minimum and maximum of total lead time of 4 configurations (1.Edge server alone, 2.Smart glasses alone, 3.802.11ngb and 4.802.11ac).

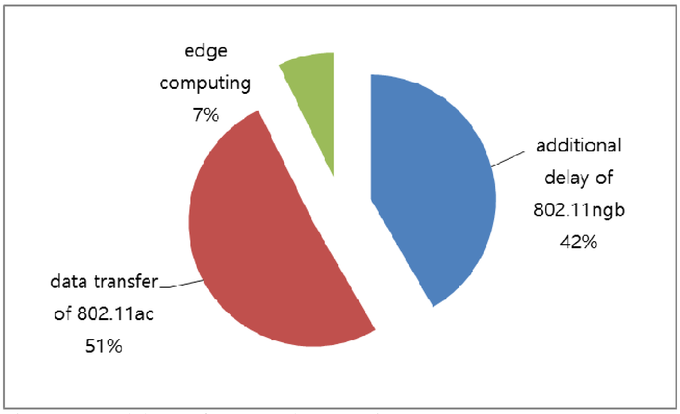


Figure 8. Breakdown of 802.11ngb connection

V. Discussion and Concluding Remarks

This paper focuses on the limitation and solution of system architecture of Smart operator supported by Wearable device such as Smart glasses. Proposed architecture of Smart glasses and Edge server distributes the functions of high-computing programs. Four different experiments compare the performances of smart glasses and edge server with two different wireless communications.

This paper shows the main contributions in computing and communication and modularity. Distribution of computing power allows the system to use various object detectors without the limitation of device platform. And as the edge server uses cabled electricity, there is no limitation of resource consumption even if high requirements of GPU/CPU. Modularity is another aspect of the advantage of proposed architecture. The vendors of smart glasses and object detectors develop their apps in different operating system. New glasses and new detector are not compliant to legacy system in some cases.

Proposed architecture use very simple data format and exchange method and allows new glasses and edge platform to integrate with existing system without re-programming. Proposed architecture will contribute the extensibility of the production system where human operator collaborate with robot and machine intensively. This production system requires for the applications smart glasses and edge server in order to support human operators in novice training, assembly guide and maintenance instruction. However Smart glasses and edge server are fast-growing market with many change of their platform or their operating system. This architecture will be the mediator to cope with the fast change of device platform and to keep existing services without additional programming.

Future developments are the application of 5G communication to provide faster connection than wired cable and new experiment of battery consumption of smart glasses in different configuration. Future research in this field will further reduce the cost of deploying these systems and become a potential conceptual concept. Future works are system engineering approaches to keep the design from factory network to specific programming function consistent so that "lean IT manufacturing system" is achieved. Future developments in this field will further reduce the cost of deploying these systems and become a potential conceptual concept. Future works are system engineering approaches to keep the design from factory network to specific programming function consistent so that "lean IT manufacturing system" is achieved.

REFERENCES

[1] D.W.F. van Krevelen and R. Poelman, "A Survey of Augmented Reality Technologies, Applications and Limitations" *The International Journal of Virtual Reality*, 2010, 9(2), pp. 1-20.

[2] V. Elia, M. Gnoni and A. Lanzilotto, "Evaluation the application of augmented reality devices in manufacturing from a process point of view: An AHP based model", *Expert Systems with Applications*, 63, pp.187-197, 2016.J. Clerk Maxwell, A Treatise on Electricity and Magnetism, 3rd ed., vol. 2. Oxford: Clarendon, 1892, pp.68–73.

[3] N. Navab, "Developing killer apps for industrial augmented reality", *IEEE Computer Graphics and Applications*, Vol. 24 (3), pp. 16-20, 2004

[4] M. Abramovici, M. Wolf, S. Adwenat and M. Neges, "Context-aware Maintenance Support for Augmented Reality Assistance and Synchronous Multi-user Collaboration", *Proceedings of the 5th International Conference on Through-life Engineering Services*, pp. 18-22, 2017.

[5] D. Antonelli and S. Astanin, "Enhancing the Quality of Manual Sport Welding through Augmented Reality Assisted Guidance. *In: Proceedings of the 9th CIRP Conference on Intelligent Computation in Manufacturing Engineering*, pp. 556-561, 2015.

[6] M.H. Stoltz, V. Giannikas, D. McFarlane, J. Strachan, J. Um and R. Srinivasan, "Augmented Reality in Warehouse Operations: Opportunities and Barriers", *IFAC PapersOnLine*, 50-1, pp. 12979-12984, 2017.

[7] D.W Mizell, "Virtual Reality and Augmented Reality in Aircraft Design and Manufacturing." *Proceedings of WESVON/94*, Anaheim, USA, 1994.

[8] M. Mura, G. Dini and F. Failli, "An integrated environment based on augmented reality and sensing device for manual assembly workstations" *Proceedings of the 48th CIRP Conference on Manufacturing Systems*, pp. 340-345, 2016.

[9] V. Paelke, "Augmented Reality in the Smart Factory: Supporting workers in an industry 4.0 environment", *2014 IEEE Emerging Technologies and Factory Automation*, Barcelona, Spain.

[10] K. Bilal and A. Erbad, "Edge Computing for Interactive Media and Video Streaming", *2017 Second International Conference on Fog and Mobile Edge Computing*, Valencia, Spain, 2017.

[11] C. Li, L. Toni, J. Zou, H. Xiong and P. Frossard, "QoE-Driven Mobile Edge Caching Placement for Adaptive Video Streaming", *IEEE Transactions on Multimedia*, PP(99), p.1, 2017.

[12] D. Zühlke, "SmartFactory–A Vision becomes Reality." *IFAC Proceedings Volumes* 42.4 (2009): 31-39.

[13] J. Huang, V. Rathod, C. Sun, M. Zhu, A. Korattikara, A. Fathi, I. Fischer, Z. Wojna, Y. Song, S. Guadarrama, and K. Murphy, "Speed/accuracy trade-offs for modern convolutional object detectors." *IEEE Computer Vision and Pattern Recognition*, Honolulu, United States, 2017.