



## Implementasi AI Pendeteksi Pola Gerak Tangan pada Game “Pong Ball” dengan Menggunakan Algoritma Convolutional Neural Network

*AI Implementations of Hand Gestures Detection in The “Pong Ball” Game Using The Convolutional Neural Network Algorithm*

Muhammad Hakam Fardana<sup>1</sup>, Muhammad Alfa Rizky Valentino<sup>2</sup> & Wahyu S. J. Putra<sup>3</sup>

<sup>1,2,3</sup>Fakultas Ilmu Komputer, Universitas Pembangunan Nasional “Veteran” Jawa Timur. Jalan Rungkut Madya Nomor 1, Gunung Anyar, Kec. Gunung Anyar, Kota SBY, Jawa Timur 60294

Email: <sup>1</sup>[20081010198@student.upnjatim.ac.id](mailto:20081010198@student.upnjatim.ac.id), <sup>2</sup>[20081010191@student.upnjatim.ac.id](mailto:20081010191@student.upnjatim.ac.id), <sup>3</sup>[wahyu.s.j.saputra.if@upnjatim.ac.id](mailto:wahyu.s.j.saputra.if@upnjatim.ac.id).

**Abstrak.** *Keyboard, mouse, dan joystick* merupakan perangkat keras yang sering digunakan untuk interaksi antara manusia dengan komputer yang mana bersifat mekanis. Meskipun akurat, tetapi model dari metode interaksi yang dilakukan tidaklah bersifat alami sebagaimana manusia berinteraksi dengan sesamanya. Untuk dapat beroperasi, diperlukan adanya kontak langsung antara pengguna dengan komputer. Hal yang diteliti disini ialah pola gerak tangan. Hal tersebut diteliti dengan menggunakan sebuah program dengan bantuan kamera *webcam* untuk melakukan pengenalan pada objek tangan. pendeteksian pola gerak tangan diimplementasikan dengan menggunakan *library cvzone* dan algoritma *Convolutional Neural Network*. Penerapan dari akurasi data program yang dibuat ialah dengan memainkan game “Pong Ball”. Uji coba menggunakan game “Pong Ball” sederhana dimana gerakannya keatas dan kebawah. Ketika tangan digerakkan keatas, maka *bat pong ball* akan keatas sesuai pola gerak yang dilakukan, begitupun sebaliknya. Dari uji coba penelitian dengan pencahayaan terang, game “Pong Ball” berjalan dengan lancar, sedangkan ketika pencahayaan yang diterima kurang atau redup, maka otomatis sensor gerak akan melambat. Tidak hanya itu, program juga mampu mendeteksi 2 sampai 4 objek tangan yang memungkinkan terjadinya permainan antar tim.

**Kata kunci:** *Deep Learning, Artificial Intelligence, Pendeteksi Pola Gerak Tangan, Pong Ball, Convolutional Neural Network, Human Computer Interactions.*

**Abstract.** *Keyboard, mouse, and joystick are hardware devices that are often used for interaction between humans and computers which are mechanical.*

*Although accurate, the model of the interaction method is not natural as humans interact with each other. To be able to operate, it is necessary to have direct contact between the user and the computer. The thing studied here is the pattern of hand movements. This was investigated by using a program with the help of a webcam camera to perform hand object recognition. hand motion pattern detection is implemented using the cvzone library and the Convolutional Neural Network algorithm. The application of the accuracy of the program data made is by playing the game "Pong Ball". Trial using a simple "Pong Ball" game where the movement is up and down. When the hand is moved up, the bat pong ball will go up according to the pattern of motion performed, and vice versa. From research trials with bright lighting, the game "Pong Ball" runs smoothly, while when the lighting received is less or dim, the motion sensor will automatically slow down. Not only that, the program is also able to detect 2 to 4 hand objects that allow the game between teams to occur.*

**Keywords:** *Deep Learning, Artificial Intelligence, Hand Gesture Detection, Pong Ball, Convolutional Neural Network, Human Computer Interactions.*

## **1 Pendahuluan**

Dengan kemajuan teknologi komputer, peran teknologi komputer yang digunakan untuk kepentingan manusia juga semakin meningkat. Salah satunya adalah bidang interaksi manusia-komputer atau IMK, juga dikenal dengan "Interaksi Manusia Komputer (HCI)". *Keyboard, Mouse, dan Joystick* adalah perangkat keras yang biasa digunakan untuk interaksi mekanis antara manusia dan komputer. Peralatannya akurat dan andal, tetapi model interaksinya tidak alami karena orang berinteraksi satu sama lain. Kita harus berhubungan langsung dengan komputer agar bisa dioperasikan. Kebutuhan akan komunikasi yang lebih natural antara pengguna dan komputer merupakan kebutuhan dari evolusi teknologi komputer itu sendiri. Ketika pengguna berinteraksi dengan lingkungan virtual (*Virtual Reality*), komputer *portable*, atau perangkat *Augmented Reality*, perangkat interaktif tradisional tidak lagi dapat diandalkan.

Teknik interaksi berbasis visi komputer adalah salah satu dari teknik interaksi alami. Teknik ini tidak mengharuskan pengguna untuk bersentuhan langsung dengan perangkat *input*. Sebaliknya, komputer menangkap dan menginterpretasikan gerakan pengguna melalui kamera video. *Input* berupa visual dapat memberikan kemampuan persepsi pada komputer seperti yang dilakukan manusia dengan mata. Tentu saja, kita tidak memerlukan *mouse* atau *keyboard* untuk berinteraksi dengan manusia. Tetapi kita perlu menggunakan tangan, mata, dan telinga kita untuk mengumpulkan berbagai macam informasi yang ada di sekitar lingkungan kita. Tangan merupakan alternatif tubuh manusia yang dapat digerakkan untuk membuat berbagai pose. Manusia secara alami terbiasa berkomunikasi dengan tangan mereka. Orang-orang menggerakkan

tangan mereka dan mengambil berbagai *gesture*. Setiap *gesture* memiliki arti dan tujuannya sendiri, sesuai dengan kesepakatan atau kesepakatan umum antara orang-orang yang berkomunikasi dengannya.

Era Revolusi Industri 4.0, sebutan untuk zaman dimana kita sebagai manusia mulai memaksimalkan penggunaan dan pengembangan teknologi yang bermanfaat untuk aktivitas sehari-hari. Kini sudah mulai banyak tren-tren teknologi yang bermunculan di zaman ini, seperti teknologi VR atau *Virtual Reality*, yakni teknologi yang dapat mereplika dunia nyata menjadi dunia digital. Dan teknologi yang paling digemari saat ini ialah AI atau *Artificial Intelligence*. Atau kita bisa sebut juga dengan Kecerdasan Buatan. Kecerdasan Buatan merupakan teknologi yang mereplika sistem-sistem yang ada pada Manusia, contohnya Otak. Sederhananya, kecerdasan buatan ialah Replika Manusia. Salah satu contoh dari Kecerdasan Buatan yang sering dijumpai yakni Robot. Robot menerapkan sistem-sistem saraf yang ada pada otak manusia. Dengan begitu, ia bisa melakukan aktivitas yang biasa dilakukan manusia.

Pengenalan pola gerak tangan sudah banyak dimanfaatkan di berbagai bidang, seperti hiburan, pendidikan, dan keamanan. Kamera atau yang biasa kita sebut dengan *webcam* adalah alat untuk mendeteksi pola gerakan pada tangan. Deteksi gerakan tangan ini diimplementasikan menggunakan *library* “cv zone”, dan salah satu metode yang paling banyak digunakan adalah “CNN” [1]. Metode ini mengasumsikan bahwa kecerahan setiap titik objek adalah konstan setiap waktu, dan piksel dihitung berdasarkan posisi vektor dan perbandingan pada *frame* berikutnya [2]. Metode ini menangkap gerakan dengan menyimpan gambar piksel dibingkai pertama dan mencari piksel terdekat dengan kecerahan yang sama dibingkai kedua. Ini akan memberi tahu kita obyek manakah yang bergerak [3]. Teknologi tersebut kemudian diimplementasikan untuk memainkan sebuah game yang bernama “Pong Ball”. Karena permainan ini merupakan permainan yang membutuhkan akurasi dan kecepatan gerak untuk memenangkan bola.

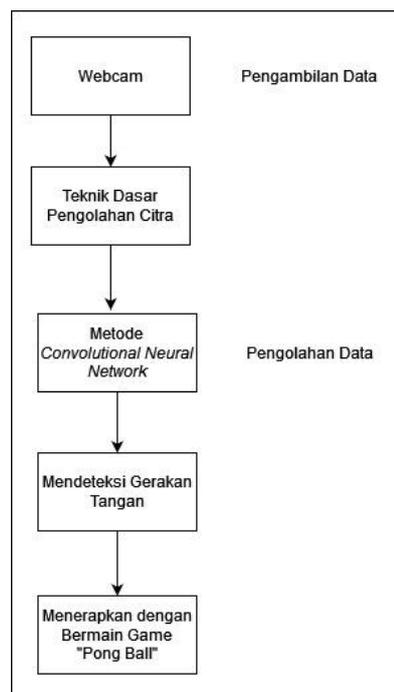
## 2 Metode penelitian

Perancangan sistem dalam penelitian ini terpecah dalam dua hal, yakni pengambilan data, dan pengolahan data. Pengambilan data merupakan hal yang paling utama untuk dilakukan dan pada langkah inilah kita mendapatkan informasi yang berguna untuk pengekseskusion sebuah masalah. Setelah pengambilan data selesai dilakukan, maka hal yang selanjutnya yakni pengolahan data yang sudah diambil dari proses pengambilan data tersebut lewat media aplikasi yang sudah dipersiapkan sebelumnya. Ilustrasinya seperti gambar berikut.



**Gambar 1.** Arsitektur Sistem

**Gambar 1** mengilustrasikan terkait bagaimana proses dari pengenalan pola gerak tangan manusia yang kemudian akan dikirimkan ke perangkat komputer. Pengenalan tersebut dilakukan melalui sebuah *webcam* sebagai tempat penerimaan data dari *user* dan kemudian diolah pada sebuah aplikasi yang nantinya dari situ akan dapat diimplementasikan pada game “Pong Ball”. Proses ini menggunakan teknik pengolahan citra digital dan metode CNN. Alur kerja sistemnya dapat diilustrasikan seperti gambar berikut.



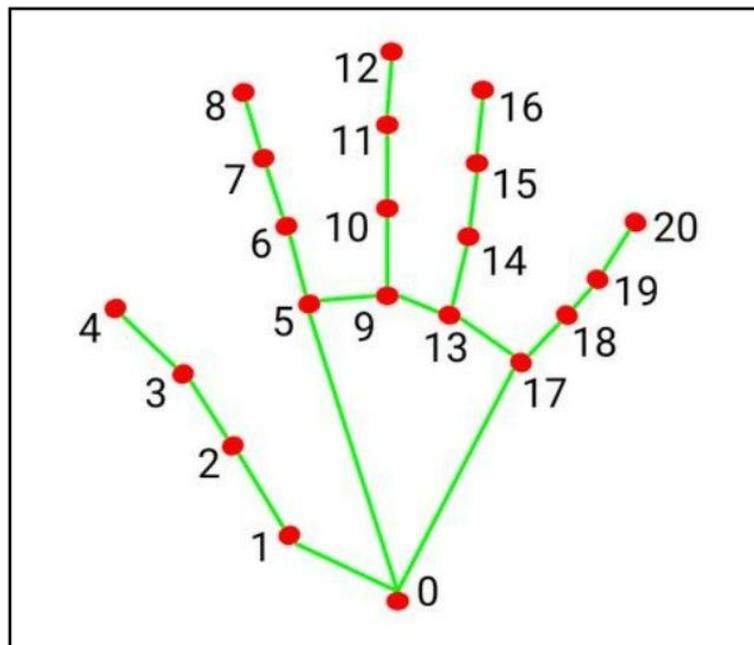
**Gambar 2.** Alur Kerja Sistem

## 2.1 Pengambilan Data (*Data Gathering*)

Pengambilan data dilakukan dengan pengoperasian oleh *webcam* pada pola gerak tangan manusia yang nantinya data tersebut akan diolah untuk proses berikutnya.

## 2.2 Pengolahan Data (*Data Processing*)

Pada proses ini, data yang sudah diperoleh sebelumnya diolah dengan menggunakan teknik pengolahan citra digital. Teknik pengolahan citra digital terdiri dari *segmentation*, *feature extraction*, dan *morfologi processing*. Setelah semua tahap dilakukan, metode CNN atau *Convolutional Neural Network* akan mendeteksi pola gerakan tangan. *Webcam* akan mendeteksi titik-titik dengan total 21 titik berdasarkan sendi dan ujung kerangka tangan manusia. Lebih



jelasnya seperti gambar berikut.

**Gambar 3.** Pengolahan Data Deteksi Gerakan Tangan

Tahap pengolahan data merupakan tahap inti dari sistem program yang akan dirancang. Tahap pengolahan data ini dibagi menjadi beberapa proses seperti yang sudah disebutkan sebelumnya, yakni sebagai berikut.

### 2.2.1 Tahap Deteksi Area

Pada tahap ini, *webcam* akan mendeteksi area dari data tangan yang akan dipakai sebagai bahan penelitian nantinya.

### 2.2.2 Tahap Segmentasi (*Segmentations*)

Pada tahap ini, proses pemisahan objek dengan latar belakang (*background*) merupakan tujuan utamanya. Sederhananya, pada tahap segmentasi adalah proses membaca *frame* pertama dan menganggapnya sebagai *frame background*. Konversi *frame background* ke *grayscale* dan *set* nilai dari *threshold*. Proses tersebut dilakukan untuk *frame* pertama, dan untuk *frame-frame* selanjutnya sampai pada tahap terakhir yakni tahap segmentasi berupa video yang sudah dikonversi menjadi sebuah gambar.

### 2.2.3 Tahap Ekstraksi Fitur (*Feature Extractions*)

Tahap selanjutnya yakni ekstraksi fitur guna untuk mendapatkan keakuratan data dari *object* di setiap *frame* gambar. Tahap dari ekstraksi dilakukan setelah tahap segmentasi. Setelah dipisahkan antara *foreground* dengan *background*, maka dicari tingkat perubahan warna dari *background* dan *foreground* [4] yang kemudian didapat data akurasi terkait *frame* gambar tersebut. Setelah tahap ekstraksi dilakukan, langkah selanjutnya yakni proses morfologi, yakni dilasi dan erosi. Inti dari proses morfologi yakni memudahkan penafsiran sebuah objek [5].

### 2.2.4 Tahap Deteksi Objek (*Object Detections*)

Tahap yang terakhir dilakukan yakni deteksi objek. Deteksi objek merupakan proses pengekstrakan fitur dari setiap objek yang ada. Berdasarkan dimensi dari setiap objek ini yang memiliki fitur khusus. Metode yang diterapkan yakni CNN. Metode ini digunakan untuk mendeteksi dan menunjukkan objek di setiap *frame*. Dalam metode ini, piksel dihitung berdasarkan pada posisi vektor dan akan dibandingkan dengan *frame* berikutnya untuk posisi piksel tersebut. Secara umum, pola gerakan yang terjadi akan sesuai dengan vektor posisi piksel [6]. Perhitungan dengan metode CNN didasarkan pada dua asumsi sebagai berikut.

1. Kecerahan dari setiap titik objek konstan setiap waktu;

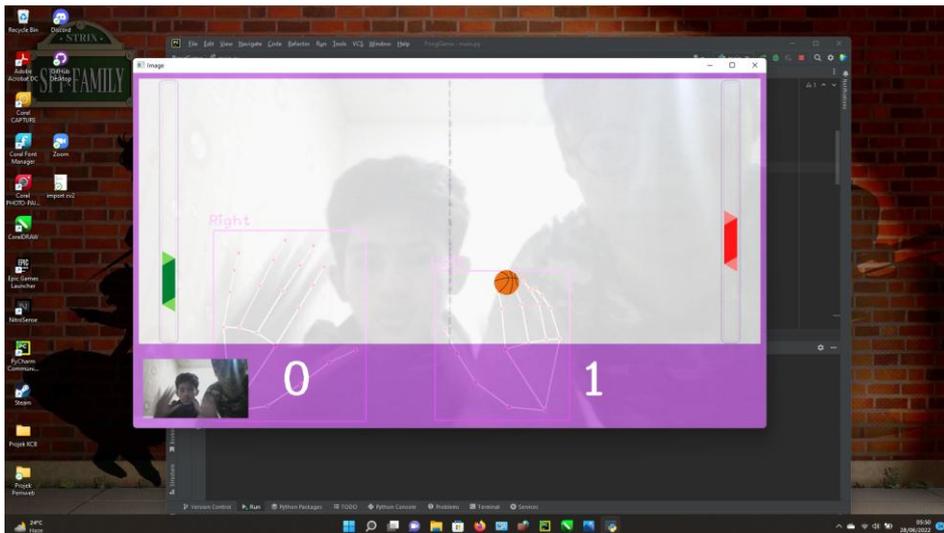
2. Pola gerakan gambar di setiap *frame* yang berubah secara berkala.

Penafsiran objek mengacu pada proses menelusuri objek yang bergerak dalam setiap *frame*. Inti dari penafsiran yakni menentukan akurasi posisi yang tepat dari suatu objek. [7].

### 3 Hasil dan Pembahasan

Hasil dari penelitian ini adalah program yang dapat mendeteksi pola gerakan tangan, baik itu tangan kanan, maupun tangan kiri. Program mendeteksi pola gerakan tangan berdasarkan pola variasi bentuk jari yang dibentuk, seperti pola angka, pola *thumb*, pola *cheese*, dan sebagainya. Program mendeteksi berbagai pola-pola tersebut yang kemudian nantinya dapat diimplementasikan pada game “Pong Ball”.

Untuk memaksimalkan fungsi dari program, dilakukan uji coba pada sebuah game yang bernama “Pong Ball” sebagai implementasi AI pendeteksi pola gerak tangan dengan menggunakan algoritma *Convolutional Neural*



*Network* (CNN). Ilustrasi pengimplementasiannya adalah sebagai berikut.

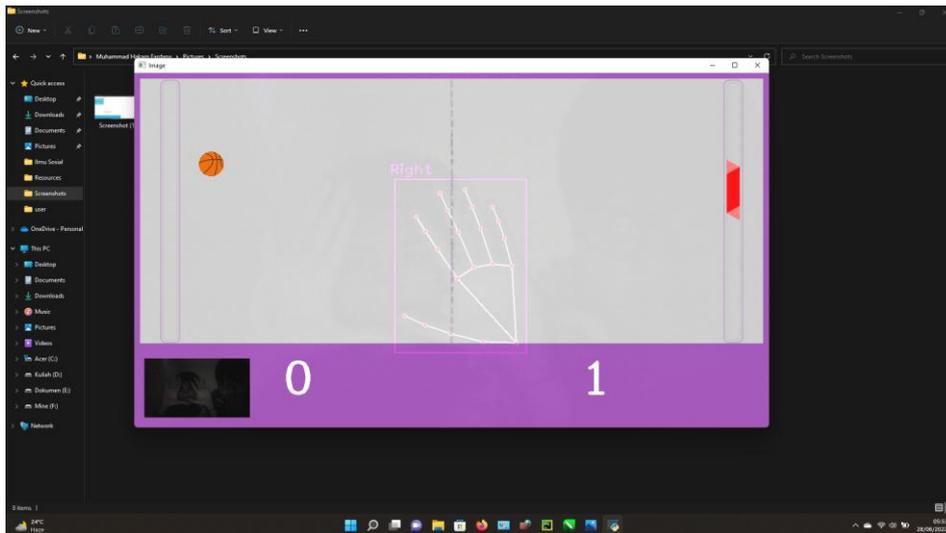
**Gambar 4.** Uji Coba dengan Game “Pong Ball”

Pada **Gambar 4** terlihat pada gambar tersebut, program mengambil data dari objek tangan kita yang kemudian diolah menjadi sebuah piksel dan

menyatukannya menjadi satu keutuhan dan kemudian membuatnya menjadi sebuah *gesture* tangan. Namun, gerakan akan semakin menurun kualitas dan keakuratan datanya jika cahaya yang diperlukan oleh *webcam* sebagai pengumpul data kurang, Maka dari itu gerakan dari programnya akan sedikit lambat.

Uji coba menggunakan game “Pong Ball” sederhana dengan gerakan keatas dan kebawah. Sebagai bantuan, untuk mendeteksi apakah pola gerakan tangan kita terdeteksi atau tidak ialah menggunakan sebuah *component* tambahan yakni sebuah *bat* atau sebuah pemukul, dalam kasus ini pemukul tersebut hanya sebatas pemukul yang tidak bergerak atau sederhananya yakni penahan objek.

Dari uji coba penelitian yang dilakukan dengan pencahayaan terang, game “Pong Ball” akan beroperasi dengan lancar saat dimainkan, sedangkan saat pencahayaan redup, sulit untuk menerima data yang berakibatkan sulit terdeteksinya pola gerak tangan yang dilakukan. Ilustrasinya seperti gambar



berikut.

**Gambar 5.** Uji Coba Pencahayaan

Terlihat pada gambar tersebut, sensor *webcam* sulit mendapatkan data yang akan diolah oleh program dikarenakan kurangnya rangsangan cahaya dari luar. Akibatnya sensor pengenalan data akan sedikit berkurang akurasiya.

Tidak hanya itu program ini juga mampu mendeteksi sampai 2-4 objek tangan yang pada pengimplementasiannya, game “Pong Ball” dapat dilakukan hingga 2 orang atau tim.

#### 4 Kesimpulan

Kesimpulan yang dapat diambil dari penelitian ini adalah.

1. Program mendeteksi pola gerak tangan untuk bermain game “Pong Ball” dengan teknik pengolahan citra dengan menggunakan metode *Convolutional Neural Network* (CNN) berjalan lancar ketika pencahayaan yang diterima terang dimana rangsangan cahaya tersebut meningkatkan keakuratan data yang diterima oleh *webcam*.
2. Jarak antara tangan dengan *webcam* juga sangat berpengaruh selain pencahayaan guna meningkatkan keakuratan data yang diterima oleh aplikasi. Tetapi, jika jarak tangan terlalu dekat dengan *webcam*, maka akan membuat program mengalami *error* karena kegagalan dalam mendeteksi objek.

#### 5 Daftar pustaka

- [1] Umar, U., Sulistyorini, R., Haryadi, A. D., 2011, Tracking Arah Gerakan Telunjuk Jari Berbasis Webcam Menggunakan Metode Optical Flow. The 13th Industrial Electronics Seminar 2011 (IES 2011) Electronic Engineering Polytechnic Institute of Surabaya (EEPIS), Surabaya, 26 Oktober 2016.
- [2] Revathi, R., Hemalatha, M., 2012, Certain Approach of Object Tracking using Optical Flow Techniques, International Journal of Computer Applications (0975 – 8887), Vol 53, No 8.
- [3] Setia Buana, I Komang, 2014, Sistem Penghitungan Jumlah Kendaraan Dengan Teknik Pengolahan Citra Berbasis Java, Konferensi Nasional Sistem dan Informatika 2014, Denpasar, 7-8 November 2014.
- [4] Kadir, A., 2008, Dasar Pemrograman Java 2, Andi Offset, Yogyakarta.
- [5] Kwong, K., Kavalier, R., Rajagopal, R., Varaiya, P., 2010, Real-Time Measurement of Link Vehicle Count and Travel Time in a Road Network, IEEE Transactions on Intelligent Transportation System, Vol 11, NO. 4.
- [6] Huang, C. H., 2011, Video-based Traffic Analysis System Using a Hierarchical Feature Point Grouping Approach, First International Conference on Robot, Vision and Signal Processing (RVSP), TBD Kaohsiung City, 21-23 November 2011.
- [7] Winarno, E., 2011, Aplikasi Deteksi Tepi pada Realtime Video menggunakan Algoritma Canny Detection. Jurnal Teknologi Informasi DINAMIK, Vol 16, No 1.
- [8] Chen, Z., Cao, J., Tang, Y., Tang, L., 2011, Tracking of Moving Objects Based on Optical Flow Detection, International Conference on Computer Science and Network Technology, 24-26 Desember 2011.

- [9] *Song, J. F., Bai, A N., Xue, R., 2011, A Reliable Counting Vehicles Method in Traffic Flow Monitoring, International Congress on Image and Signal Processing, 15-17 Oktober 2011.*
- [10] Munir Oudah, Ali Al-Naji and Javaan Chahl, " Review Hand Gesture Recognition Based on Computer Vision: A Review of Techniques" 23 July 2020.
- [11] Rafiqul Zaman Khan and Noor Adnan Ibraheem "HAND GESTURE RECOGNITION: A LITERATURE REVIEW" July, 2021.
- [12] M. Wegmuller, J. P. von der Weid, P. Oberson, dan N. Gisin, "High Resolution fiber distributed measurements with coherent OFDR," Proc. ECOC 00, 2000, paper 11.3.4, hal. 109.
- [13] Hambling, B. and Van Goethem, P., 2013, May. User acceptance testing: a step-by-step guide. Chippenham: BCS.
- [14] Lewis, J.R., 2006. Usability testing. Handbook of human factors and ergonomics, 12, p.e30.