**Raport stiintific**

**de cercetare-dezvoltare in cadrul Cloudifier SRL**

|  |  |
| --- | --- |
| **Nume proiect** | Platforma de migrare automatizată în cloud a aplicațiilor și sistemelor informatice clasice cloudifier.net |
| **Beneficiar** | CLOUDIFIER SRL |
| **Cod MySMIS** | 104349 |
| **Nr. iregistrare** | P\_38\_543 |
| **Director Proiect** | Andrei Ionut DAMIAN |
| **Activitate conform planului de proiect** | 1. Activităţi de cercetare-dezvoltare (cercetare industrială şi/sau dezvoltare experimentală) - 1.1 State-of-the-art |
| **Echipa de cercetare-dezvoltare** | Andrei Ionut DAMIAN  Octavian BULIE |
| **Livrabil** | Raport State-of-the-Art |

Contents

[1 Abstract 3](#_Toc470772468)

[2 Analiza contextului cercetarii 3](#_Toc470772469)

[2.1 Acronime 3](#_Toc470772470)

[2.2 Obiectivele proiectului 4](#_Toc470772471)

[2.3 Contextul proiectului 5](#_Toc470772472)

[2.4 Sumarul activitatii de cercetare si rezultatele 7](#_Toc470772473)

[3 Rezultatele cercetarii 9](#_Toc470772474)

[3.1 Problematica concreta analizata 9](#_Toc470772475)

[3.2 Ecosistemul Cloudifier.NET ver 1.0 12](#_Toc470772476)

[4 Anexa – Rapoarte lunare 13](#_Toc470772477)

[4.1 Raport stiintific lunar 1 13](#_Toc470772478)

[4.2 Raport stiintific lunar 2 17](#_Toc470772479)

[4.3 Raport stiintific lunar 3 20](#_Toc470772480)

# Abstract

In prezentul document este descris rezultatul etapei de cercetare industriala aferenta Activitatii 1 de cercetare industrială şi/sau dezvoltare experimentală, sub-activitatea “State-of-the-Art”. Cercetarea industriala realizata in aceasta etapa a constat in revizuirea stadiului actual al tehnicii/tehnologiilor atat in domeniul Cloud Computing / virtual desktop dar mai ales in domeniul recunoasterii si segmentarii semantice a imaginilor prin algoritmi avansati de inteligenta artificiala, domeniu de cercetare ce a luat o deosebita amploare in ultimile 12 luni.

Scopul analizei, recunoasterii, segmentarii semantice si a constructiei unei harti a imaginilor este de incorporare a acestor tehnologii avansate de inteligenta artificiala (Machine Learning) in motorul de migrare automatizata a aplicatiilor in mediul cloud. Intuitiv acest mecanism functioneaza similar unui actor/observator uman care analizeaza o imagine dupa care reproduce intr-un alt mediu din memorie elementele principale ale imaginii respective impreuna cu pozitia, forma si functionalitate lor in contextul analizat initial.

# Analiza contextului cercetarii

## Acronime

**Machine Learning** = domeniu de cercetare-dezvoltare din cadrul Inteligentei Artificiale, aflat la fundamentul acesteia, prin care un sistem computerizat este capabil sa invete automatizat din propriile greseli cu sau fara supervizare sau orice forma de factor uman

**Desktop computing** = mod de lucru in care aplicatiile utilizate sunt rulate exclusiv pe o statie de lucru sau laptop fara a exista acces la o baza de date centralizata sau la sisteme online

**Web- based computing** = mod de lucru in care aplicatiile sunt rulate in cadrul unui server de web iar utilizatorii acceseaza functionalitatile respective prin intermediul browser-elor de internet

**Legacy** = Aplicatie sau modul dezvoltat in tehnologii invechite si/sau perimate care ruleaza in mediu de tip desktop computing

**Client-Server** = mod de lucru cu masive de date prin care acestea sunt centralizate pe un server de baze de date relationale

**Segmentare semantica** = analiza unei multimi, imagini, etc si identificarea elementelor componente ale acesteia

**Cloud computing** = tehnologie si mediu de rulare a sistemelor informatice prin care toate datele si functionalitatile aplicatiilor se acceseaza prin intermediului internetului si a aplicatiilor de tip web-based

**Bot** = aplicatie care simuleaza un interlocutor cu care utilizatorul comunica in limbaj natural (roBot) pentru rezolvarea unui set de probleme analizate cu ajutorul algortmilor de tip Machine Learning

## Obiectivele proiectului

In conformitate cu Cererea de finantare aferenta contractului de finantare nr 98/09.09.2016 obiectivul proiectului „PLATFORMA DE MIGRARE AUTOMATIZATA IN CLOUD A APLICATIILOR SI SISTEMELOR INFORMATICE CLASICE- Cloudifier.NET” este cercetarea, dezvoltarea si punerea in functiune in mediul comercial a produsului platforma inovativ Cloudifier.NET, ce se adreseaza domeniului tehnologiilor informatiei si comunicatiilor. In cadrul acestui obiectiv mentionam si intentia de diseminare publica partiala a rezultatelor proiectului sub licenta European Public License.

Obiectivul proiectului “PLATFORMA DE MIGRARE AUTOMATIZATA IN CLOUD A APLICATIILOR SI SISTEMELOR INFORMATICE CLASICE - Cloudifier.NET” raspunde uneia dintre prioritatile stabilite de Comisia Europeană prin Agenda Digitala 2020 si anume stimularea si facilitarea comertului electronic la nivelul tarilor membre. Acest proiect va contribui major la sporirea utilizarii, calitatii si a accesului la tehnologia informatiei si comunicatiilor atat la nivelul local al Romaniei cat si la nivelul comunitatii de viitori potentiali utilizatori din intreaga Uniune Europeana iar implementarea cu succes a platformei va duce la sporirea contributiei sectorului TIC pentru competitivitatea economica a Romaniei.

De asemenea, proiectul are ca obiectiv integrarea federalizarii cu conceptul de virtual desktop online, oferind astfel utilizatorului posibilitatea de a integra toate aplicatiile pe care le utilizeaza in mediul online – aplicatii de tip Cloud– intr-un singur spatiu virtual, in care sa dispuna de servicii de securizare si confidentialitate avansata a informatiilor.

La finalizarea implementarii produsul final va functiona ca o platforma de tip desktop online, bazata pe tehnologii de tip Cloud Computing, cu functionalitati multi-scop, care va permite sustinerea proiectelor inovative din domeniul tehnologiei informatiilor si comunicatiilor si va facilita dezvoltarea sustenabila a acestora.

## Contextul proiectului

Produsul Cloudifier.NET revendica prin unicitatea conceptului sau o serie de functionalitati total inovative, ce nu se regasesc la produsele comerciale deja existente in piata. In procesul de analiza preliminara a stadiului actual al tehnologiei (state-of-the-art), s-au luat in considerare o serie de produse comerciale relativ similare, pentru a se face o paralela intre acestea si Cloudifier.NET si a se determina detaliile avansului tehnologic dincolo de stadiul actual al tehnologiei (advances beyond state-of-the-art).

Implementarea proiectului „Platforma de migrare automatizata in Cloud a aplicatiilor si sistemelor informatice clasice- Cloudifier.NET” va permite:

* Sustinerea proiectelor inovative din domeniul tehnologiei informatiilor si comunicatiilor;
* Va facilita dezvoltarea sustenabila a acestora, prin oferirea accesului la o platforma de tip comunitate, in care sa poata gasi si regasi aplicatii si sisteme la cerere, utilizand astfel exclusiv structura de costuri de tip OPEX, fata de structurile clasice de tip CAPEX;
* Sprijinirea companiilor mari prin servicii “cost-effective” pentru migrarea de la aplicatii legacy- (aplicatii dezvoltate prin metode clasice de programare si implementare de tip desktop sau client-server, ce utilizeaza resurse bazate pe cheltuieli de capital, cum ar fi echipamente de calcul locale, licente locale, s.a.m.d - spre deosebire de modelul de aplicatii bazat pe tehnologia Cloud Computing care utilizeaza resurse la cerere scalabile, elastice si bazate aproape exclusiv pe modelul de cheltuieli operationale OPEX)" - la aplicatii in Cloud;
* Sprijinirea atat a mediului IMM, cat si a utilizatorilor privati, in vederea accesului la un mediu de tip spatiu virtual, personal de lucru online, in continua dezvoltare.

Unificarea avansului tehnologic adus de subplatforma avansata de translatare bazata de Machine Learning a aplicatiilor clasice in medii de tip Cloud cu avansul tehnologic propus de subplatforma de federalizare, brokeraj de date si spatiu virtual de lucru (virtual desktop), face ca platforma sa ofere o arie de inovare deosebit de generoasa, acoperind mai multe nevoi orizontale in domeniul tehnologiilor informatiilor si comunicatiilor, ce au impact asupra unor arii multiple tehnologice si variate industrii.

Cele doua subsisteme principale ale platformei Cloudfier sunt:

1. subsistemul bazat pe Machine Learning de translatare inteligenta automatizata a aplicatiilor clasice desktop in aplicatii online in mediu de tip Cloud computing cu accent pe trecerea de la modele de sisteme informatice bazate pe CAPEX la modele de sisteme informatice bazate pe OPEX;
2. subsistemul de federalizare a aplicatiilor de tip Cloud Computing provenite din surse multiple in vederea realizarii unei platforme de tip spatiu de lucru personal virtual online (online personal virtual desktop).

In concluzie directiile principale de utilitate ale platformei propuse sunt axate pe trei mari categorii distribuite in doua zone de inovatie dupa cum urmeaza:

1. Inovare in domeniu federalizarii platformelor, migrarii datelor si spatiilor personale virtuale:
   1. broker de servicii de Cloud pe care se inregistreaza furnizorii de servicii noi si inovative de Cloud;
   2. agregator si federalizator de servicii de Cloud prestandardizate
2. Inovare in domeniul migrarii aplicatiilor construite pe principiile clasice ale sistemelor informatice catre noile paradigme tehnologice definite de Cloud Computing:
   1. sistem inteligent bazat pe tehnici avansate de Machine Learning destinat traducerii applicatiilor clasice desktop sau aplicatiilor client-server in aplicatii de tip Cloud Computing;
   2. provizionarea automatizata a aplicatiilor translatate in mediul platformei inovative

## Sumarul activitatii de cercetare si rezultatele

Activitatea de cercetare a stadiului actual al tehnologiei s-a axact atat pe o serie de proiecte depuse in cadrul apelurilor de proiecte de cercetare beyond-state-of-the-art din cadrul Horizon 2020 dar mai ales pe lucrari publicate in cele mai renumite jurnale stiintifice cum ar fi:

* Journal of Machine Learning Research, www.jmlr.org (ISSN 1533-7928)
* International Journal of Neural Systems, ISSN:0129-0657
* IEEE Transactions on Pattern Analysis and Machine Intelligence, ISSN:0162-8828
* IEEE Transactions on Neural Networks and Learning Systems, ISSN:2162-237X
* Machine Learning, ISSN:0885-612
* International Journal of Intelligent Systems, ISSN:0884-8173
* Expert Systems with Applications, ISSN:0957-4174

De mentionat este faptul ca au fost analizate in mod particular lucrari ale celor mai cunoscuti cercetatori din domeniul inteligentei artificiale ca:

* Yoshua Bengio: http://www.iro.umontreal.ca/~ben...
* Geoffrey Hinton: http://www.cs.toronto.edu/~hinton/
* Alex Smola: http://alex.smola.org/
* Andrew Ng: http://ai.stanford.edu/~ang/
* Alex Krizhevsky: http://www.cs.toronto.edu/~kriz/
* Ilya Sutskever: http://www.cs.toronto.edu/~ilya/
* Andrej Karpathy: http://cs.stanford.edu/people/karpathy/
* Francois Chollet: https://blog.keras.io/author/francois-chollet.html
* Chris Meek: http://research.microsoft.com/en...
* Hugo Larochelle: http://www.dmi.usherb.ca/~larocheh/index\_en.html
* Ian Goodfellow: https://en.wikipedia.org/wiki/Ian\_Goodfellow
* Aaron Courville: https://aaroncourville.wordpress.com/
* Leo Breiman: http://www.stat.berkeley.edu/~br...
* Andrew McCallum: http://www.cs.umass.edu/~mccallum/
* Chris Meek: <http://research.microsoft.com/en>...
* Trevor Darrell: https://people.eecs.berkeley.edu/~trevor/
* Jonathan Long: http://people.eecs.berkeley.edu/~jonlong/

Din lucrarile analizate au fost extrase si analizate elemente cum ar fi:

1. Modelele de reconstructie arhitecturala si regenerare cod sursa
   1. Modele si metode de analiza a segmentarii ecranelor din cadrul aplicatiilor legacy bazate pe algoritmi clasici interativi si modele de Machine Learning fara retele neurale adanci
   2. Modele si metode de segmentare semantica cu ajutorul retelelor neurale adanci convolutionale a componentelor din cadrul ecranelor aplicatiilor legacy
2. Modelele de migrare aplicatie
   1. Utilizarea retelelor neurale recurente si in particular a LSTM (Long Short Term Memory) pentru realizarea de sisteme expert de tip bot care sa fie capabile de a asista utilizatorul in procesul de dezasamblare in subcomponente a aplicatiilor legacy si reasamblarea automatizata a acestora in medii de tip Cloud computing

Deasemenea in vederea atingerii tuturor elementelor de analiza propuse au fost analizate tehnologii si abordari pentru probleme de inginerie si dezvoltare experimentala cu complexitate stiintifica mai redusa cum ar fi:

1. Modelele de analiza ale tehnologiei de tip baza de date utilizata in aplicatia legacy ce urmeaza a fi migrata
   1. Baze de date flat-file
   2. Baze de date relationale nebazate pe tehnologii client-server
2. Modelele de analiza ale entitatilor de date
   1. Analiza bazata pe parcurgerea si recunoasterea grafurilor
3. Modelele de analiza ale relationarii intre entitatile de date
   1. Analiza bazata pe parcurgerea si recunoasterea grafurilor
4. Modelele de recunoastere a tehnologiei componentelor/modulelor sistemului
   1. Analiza bazata pe citirea componentelor sistemului de operare gazda
5. Modelele de recunoasterele ale layere-lor de comunicare/API/etc.
   1. Analiza bazata pe citirea componentelor sistemului de operare gazda

# Rezultatele cercetarii

## Problematica concreta analizata

Scopul analizei stadiului actual al tehnologiei este acela de a determina metodele cele mai moderne/actuale de realizare a predictiilor/inferentelor in imagistica – in particular in cazul proiectului CLOUDIFIER referindu-ne la analiza imaginilor captate in timp real in timpul functionarii aplicatiilor si implicit analiza automatizata cu ajutorul recunoasterii avansate de forme/imagini a aplicatiilor “legacy” in vederea transalatarii acestora automatizate.

In decursul lunii octombrie 2016 au fost analizate cele mai recente si avansate lucrari de cercetare fundamentala si industriala provenite de la cele mai prestigioase institute si universitati printre care enumeram:

* Caltech – California Institute for Technology
* MIT – Massachusetts Institute for Technology
* Stanford
* University of Toronto
* Harvard
* University of Washington

Principalele zone analizate au fost:

* Metodele de tip Deep Learning bazate pe Retele Neuronale Convolutionale – Deep Convolutional Neural Networks
* Metode de tip shallow learning pentru invatarea supervizata a structurilor si a elementelor de imagistica utilizand modele de invatare in timp real (online learning)
* Cele mai moderne abordari in Deep Learning – Tensor Flow
* Cele mai moderne abordari in shallow learning – Extreme Boosted Decision Trees / Random Forests - XGBoost

In decursul lunii noiembrie 2016 au fost analizate cele mai recente si avansate lucrari de cercetare fundamentala in vederea determinarii unui set de algoritmi de Machine Learning ideali pentru identificarea primitivelor de interfata grafica (butoane, campuri, ferestre, texte statice, etc) si a pozitiei acestora in cadrul ecranelor interfetelor grafice

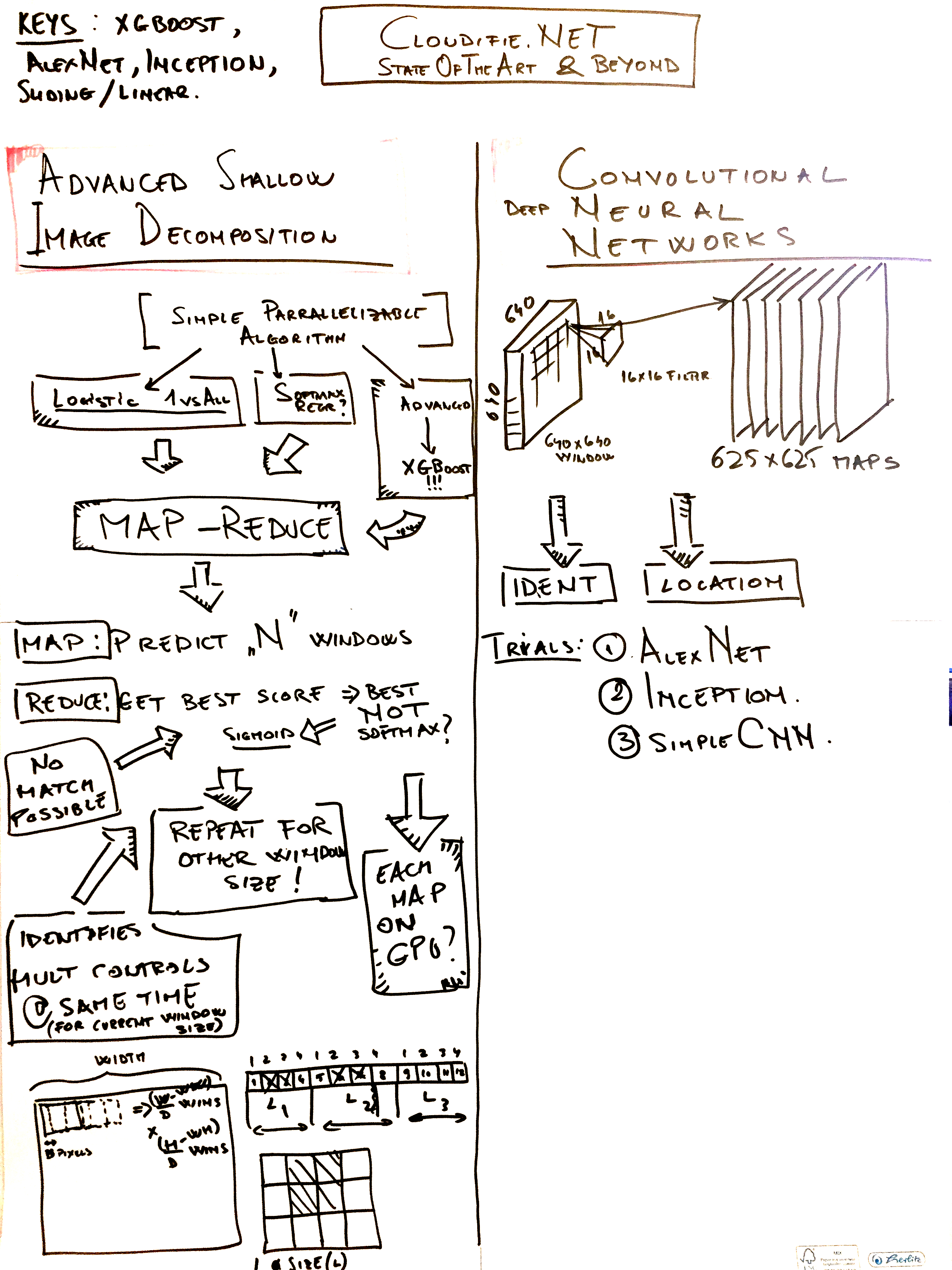
* Determinarea unui algoritm de tip Machine Learning pentru generarea AUTOMATA de interfe grafice si cod sursa aferente pe baza schitelor facute manual pe suport de hartie, tabla, etc
* Analiza TensorFlow
* Analiza XGBoost
* Analiza metode si propuneri pentru biblioteci interne
* Analiza si testarea experimentala a mediilor de procesare numerica masiv paralela cu ajutorul GPU (tehnologiile bazate pe nuclee de calul masiv paralel CUDA)

In decursul luni decembrie 2016 a fost continuat procesul de analiza a stadiului curent al tehnologiei in domeniul sistemelor de tip Machine Learning cu accent pe zona de Deep Learning si in particular a sistemelor de analiza si recunoastere bazata pe inteligenta artificiala a imaginilor. In decursul acestei luni analiza stadiului curent al cercetarii fost fost axat in principal pe lucrarea stiintifica publicata recent de J. Long et al “Fully Convolutional Networks for Semantic Segmentation”, lucrare considerata actualmente state-of-the-art in ceea priveste metodele de recunoastere si segmentare a componentelor in cadrul imaginilor. Pentru referinta prezentam anexat un scurt rezumat in limba engleza a lucrarii de referinta.

Principalele puncte pe care le urmarim in cercetare sunt urmatoarele:

1. Determinarea metodelor optime bazate pe Deep Learning pentru recunoasterea si segmentarea (identificarea locatiei spatiale) a elementelor de interfata grafica pe care Cloudifier.NET va trebuie sa le translateze automatizat din aplicatiile legacy in aplicatiile din mediul cloud computing.
2. Aplicarea de metode simple bazate pe algoritmi de machine learning superficiali (regresie logistica, arbori de decizie, clasificare naiva bazata pe teorema lui Bayes, clusterizare cu analiza distantelor euclidiene) precum si metode de segmentare iterativa a imaginilor analizate cum ar fi metoda ferestrelor deplasate continuu (“ferestre alunecatoare” sau sliding-windows algorithm)

In Imaginea nr. 1 (“Schita de principiu a abordarii stadiului actual al tehnologiei si avansurile aferente”) este prezentata abordarea celor doua metode care vor fi proiectate in paralel in procesul de proiectare a modelelor arhitecturale ale etapei 1.2 din activitatile proiectului si ulterior experimentate in cadrul procesului de dezvoltare experimentala. Detaliile acestei abordari vor fi prezentate in sectiunea “3.3 Doua directii de abordare avansata a cercetarii”

***Imaginea 1 “Schita de principiu a abordarii stadiului actual al tehnologiei si avansurile aferente”***

## Ecosistemul Cloudifier.NET

***Diagrama 2 – Ecosistemul Cloudifier.NET ver 1.0***

|  |
| --- |
| **State-of-the-art** |
| **Models for data structure analysis and inference**  **Models for user interface**  **Models for stack analysis**  **Models for automated migration** |

|  |  |  |
| --- | --- | --- |
| **Heuristic approach for application migration based on Artificial Intelligence Machine Learning** | | |
| **Cloudifier.NET CAT**  **Automated Translation Engine Architecture** | **Cloudifier.NET Virtual Desktop Renderer** | **Experimental ML Models** |
| **Cloud environment** | | |

Ecosistemul Cloudifier.NET descris in Diagrama 2

## Doua directii de abordare avansata a cercetarii

Conform diagramei descrise anterior in urma analizei stadiului actual al tehnologiei au fost determinate doua directii principale in care se vor desfasura activitati de modelare arhitecturala si dezvoltare experiementala in urmatoarea perioada:

1. Modele avansate de recunoastere si decompoziei a cadrelor (imaginilor) din aplicatiile legacy prin utilizarea de algoritmi paralelizabili de predictie bazata pe tehnici de machine learning fara retele neurale adanci
2. Modele de segmentare semantica precum si proiectare de sisteme automatizate expert de tip “Bot” cu ajutorul retelelor neurale adanci convolutionale si recurente (LSTM)

# Anexa – Rapoarte lunare

## Raport stiintific lunar 1

**Raport stiintific**

**de cercetare-dezvoltare in cadrul Cloudifier SRL**

**Nr. 62/31.10.2016**

|  |  |
| --- | --- |
| **Nume proiect** | Platforma de migrare automatizată în cloud a aplicațiilor și sistemelor informatice clasice cloudifier.net |
| **Beneficiar** | CLOUDIFIER SRL |
| **Cod MySMIS** | 104349 |
| **Nr. iregistrare** | P\_38\_543 |
| **Director Proiect** | Andrei Ionut DAMIAN |
| **Activitate conform planului de proiect** | 1. Activităţi de cercetare-dezvoltare (cercetare industrială şi/sau dezvoltare experimentală) - 1.1 State-of-the-art |
| **Luna** | Octombrie 2016 |
| **Echipa de cercetare-dezvoltare** | Andrei Ionut DAMIAN  Octavian BULIE |
| **Descrierea activitatilor desfasurate activitatii** | In decursul acestei luni a inceput procesul de analiza a stadiului curent al tehnologiei in domeniul sistemelor de tip Machine Learning cu accent pe zona de Deep Learning, domeniu de cercetare in dezvoltare la nivel international ce a luat amploare deosebita in ultimii 5 ani.  Scopul analizei stadiului actual al tehnologiei este acela de a determina metodele cele mai moderne/actuale de realizare a predictiilor/inferentelor in imagistica – in particular in cazul proiectului CLOUDIFIER referindu-ne la analiza imaginilor captate in timp real in timpul functionarii aplicatiilor si implicit analiza automatizata cu ajutorul recunoasterii avansate de forme/imagini a aplicatiilor “legacy” in vederea transalatarii acestora automatizate.  In decursul lunii octombrie 2016 au fost analizate cele mai recente si avansate lucrari de cercetare fundamentala si industriala provenite de la cele mai prestigioase institute si universitati printre care enumeram:   * Caltech – California Institute for Technology * MIT – Massachusetts Institute for Technology * Stanford * University of Toronto * Harvard * University of Washington   Principalele zone analizate au fost:   * Metodele de tip Deep Learning bazate pe Retele Neuronale Convolutionale – Deep Convolutional Neural Networks * Metode de tip shallow learning pentru invatarea supervizata a structurilor si a elementelor de imagistica utilizand modele de invatare in timp real (online learning) * Cele mai moderne abordari in Deep Learning – Tensor Flow * Cele mai moderne abordari in shallow learning – Extreme Boosted Decision Trees / Random Forests - XGBoost   In decursul lunilor noiembrie si decembrie se va continua analiza conform graficului de implementare a proiectului cu accent pe urmatoarele:   * Determinarea unui algoritm ideal pentru identificarea primitivelor de interfata grafica (butoane, campuri, ferestre, texte statice, etc) si a pozitiei acestora in cadrul ecranelor interfetelor grafice * Determinarea unui algoritm de tip Machine Learning pentru generarea AUTOMATA de interfe grafice si cod sursa aferente pe baza schitelor facute manual pe suport de hartie, tabla, etc * Analiza TensorFlow * Analiza XGBoost * Analiza metode si propuneri pentru biblioteci interne * Analiza si testarea experimentala a mediilor de procesare numerica masiv paralela cu ajutorul GPU (tehnologiile bazate pe nuclee de calul masiv paralel CUDA) |

|  |  |  |
| --- | --- | --- |
| Perioada | Efort in ore-om | Descriere |
| 3.10.2016-  14.10.2016 | 160 | Selectia si analiza preliminara a celor mai importante lucrcari din domeniul recunoasterii de imagini cu ajutorul retelelor adanci neurale convolutionale (Deep Convolutional Neural Networks). A fost inceputa analiza state-of-the-art pe ultimile cercetari realizate de laboratoarele de cercetare ale Google in Inteligenta Artificiala – biblioteca TensorFlow |
| 17.10.2016-  21.10.2016 | 80 | Analiza XGBoost – actualmente cea mai puternica infrastructura si biblioteca de shallow learning bazata pe modele de tip ansamblu |
| 24.10.2016-  31.10.2016 | 96 | Inceperea efectuarii de teste experimentale pe modele arhitecturale simple bazate pe regresii logistice adaptate si optimizate online si retele neural cu conectare completa.  Testele s-au realizat dupa cum urmeaza:   * Python cu ajutorul:   + Sci-Kit-Learn   + Biblioteca dezvoltata intern in cadrul Cloudifier pentru regresii logistice avansate (OnlineClassifierEngine.py)   + Biblioteca de retele neurale cu conectivitate completa realizata in cadrul Cloudifier * C++ cu ajutorul bibliotecii de calcul numeric optimizat Eigen |

## Raport stiintific lunar 2

**Raport stiintific**

**de cercetare-dezvoltare in cadrul Cloudifier SRL**

**Nr. 98/29.12.2016**

|  |  |
| --- | --- |
| **Nume proiect** | Platforma de migrare automatizată în cloud a aplicațiilor și sistemelor informatice clasice cloudifier.net |
| **Beneficiar** | CLOUDIFIER SRL |
| **Cod MySMIS** | 104349 |
| **Nr. iregistrare** | P\_38\_543 |
| **Director Proiect** | Andrei Ionut DAMIAN |
| **Activitate conform planului de proiect** | 1. Activităţi de cercetare-dezvoltare (cercetare industrială şi/sau dezvoltare experimentală) - 1.1 State-of-the-art |
| **Luna** | Noiembrie 2016 |
| **Echipa de cercetare-dezvoltare** | Andrei Ionut DAMIAN  Octavian BULIE |
| **Descrierea activitatilor desfasurate activitatii** | In decursul acestei luni a fost continuat procesul de analiza a stadiului curent al tehnologiei in domeniul sistemelor de tip Machine Learning cu accent pe zona de Deep Learning, domeniu de cercetare in dezvoltare la nivel international ce a luat amploare deosebita in ultimii 5 ani. Scopul analizei stadiului actual al tehnologiei este acela de a determina metodele cele mai moderne/actuale de realizare a predictiilor/inferentelor in imagistica – in particular in cazul proiectului CLOUDIFIER referindu-ne la analiza imaginilor captate in timp real in timpul functionarii aplicatiilor si implicit analiza automatizata cu ajutorul recunoasterii avansate de forme/imagini a aplicatiilor “legacy” in vederea transalatarii acestora automatizate.  In decursul lunii noiembrie 2016 au fost analizate cele mai recente si avansate lucrari de cercetare fundamentala in vederea determinarii unui set de algoritmi de Machine Learning ideali pentru identificarea primitivelor de interfata grafica (butoane, campuri, ferestre, texte statice, etc) si a pozitiei acestora in cadrul ecranelor interfetelor grafice   * Determinarea unui algoritm de tip Machine Learning pentru generarea AUTOMATA de interfe grafice si cod sursa aferente pe baza schitelor facute manual pe suport de hartie, tabla, etc * Analiza TensorFlow * Analiza XGBoost * Analiza metode si propuneri pentru biblioteci interne * Analiza si testarea experimentala a mediilor de procesare numerica masiv paralela cu ajutorul GPU (tehnologiile bazate pe nuclee de calul masiv paralel CUDA)   In decursul lunii decembrie se va continua analiza inceputa in lunile octombrie si noiembrie conform graficului de implementare a proiectului. |

| **Perioada** | **Efort in ore-om** | **Descriere** |
| --- | --- | --- |
| 01.11.2016-  11.11.2016 | 144 | Analiza model arhitectural de Retea Neurala total conectata in Python pentru analiza elementelor de interfata grafica si testarea experimentala a acesteia pe seturi de data de imagistica. Analiza PyCUDA – infrastructura programabila pentru procesoarele de calcul numeric masiv paralel CUDA |
| 14.11.2016-  30.11.2016 | 208 | Testarea in regim CPU si GPU a celui mai recent framework de Deep Learning realizat de laboaratoarele de cercetare ale Google – TensorFlow |

## Raport stiintific lunar 3

**Raport stiintific lunar**

**de cercetare-dezvoltare in cadrul Cloudifier SRL**

**Nr. 112/9.12.2016**

|  |  |
| --- | --- |
| **Nume proiect** | Platforma de migrare automatizată în cloud a aplicațiilor și sistemelor informatice clasice cloudifier.net |
| **Beneficiar** | CLOUDIFIER SRL |
| **Cod MySMIS** | 104349 |
| **Nr. iregistrare** | P\_38\_543 |
| **Director Proiect** | Andrei Ionut DAMIAN |
| **Activitate conform planului de proiect** | 1. Activităţi de cercetare-dezvoltare (cercetare industrială şi/sau dezvoltare experimentală) - 1.1 State-of-the-art |
| **Luna** | decembrie 2016 |
| **Echipa de cercetare-dezvoltare** | Andrei Ionut DAMIAN  Octavian BULIE |
| **Descrierea activitatilor desfasurate activitatii** | In decursul acestei luni a fost continuat procesul de analiza a stadiului curent al tehnologiei in domeniul sistemelor de tip Machine Learning cu accent pe zona de Deep Learning si in particular a sistemelor de analiza si recunoastere bazata pe inteligenta artificiala a imaginilor.  In decursul acestei luni analiza stadiului curent al cercetarii fost fost axat in principal pe lucrarea stiintifica publicata recent de J. Long et al “Fully Convolutional Networks for Semantic Segmentation”, lucrare considerata actualmente state-of-the-art in ceea priveste metodele de recunoastere si segmentare a componentelor in cadrul imaginilor. Pentru referinta prezentam anexat un scurt rezumat in limba engleza a lucrarii de referinta.  Principalele puncte pe care le urmarim in cercetare sunt urmatoarele:   1. Determinarea metodelor optime bazate pe Deep Learning pentru recunoasterea si segmentarea (identificarea locatiei spatiale) a elementelor de interfata grafica pe care Cloudifier.NET va trebuie sa le translateze automatizat din aplicatiile legacy in aplicatiile din mediul cloud computing. 2. Aplicarea de metode simple bazate pe algoritmi de machine learning superficiali (regresie logistica, arbori de decizie, clasificare naiva bazata pe teorema lui Bayes, clusterizare cu analiza distantelor euclidiene) precum si metode de segmentare iterativa a imaginilor analizate cum ar fi metoda ferestrelor deplasate continuu (“ferestre alunecatoare” sau sliding-windows algorithm) |

| **Perioada** | **Efort in ore-om** | **Descriere** |
| --- | --- | --- |
| 01.12.2016-  09.12.2016 | 96 | Continuarea analizei metodelor de recunoastere a imaginilor prin CNN (Convolutional Deep Neural Networks) |
| 12.12.2016-  31.12.2016 | 224 | Analiza unui model arhitectural Alpha ce urmeaza a fi definitivat in cadrul activitatii 1.2 de cercetare. Sistemul/model arhitectural Alpha va consta in construirea unui model matematic predictiv care sa poate recunoaste elemente simple de interfata grafica de utilizator (meniu, buton, etc) si sa poata reda locatia si 1-2 alte atribute de baza ale acestora |

**Fully Convolutional Networks for Semantic Segmentation**

Evan Shelhamer, Jonathan Long, Trevor Darrell

(Submitted on 20 May 2016)

Convolutional networks are powerful visual models that yield hierarchies of features. We show that convolutional networks by themselves, trained end-to-end, pixels-to-pixels, improve on the previous best result in semantic segmentation. Our key insight is to build "fully convolutional" networks that take input of arbitrary size and produce correspondingly-sized output with efficient inference and learning. We define and detail the space of fully convolutional networks, explain their application to spatially dense prediction tasks, and draw connections to prior models. We adapt contemporary classification networks (AlexNet, the VGG net, and GoogLeNet) into fully convolutional networks and transfer their learned representations by fine-tuning to the segmentation task. We then define a skip architecture that combines semantic information from a deep, coarse layer with appearance information from a shallow, fine layer to produce accurate and detailed segmentations. Our fully convolutional network achieves improved segmentation of PASCAL VOC (30% relative improvement to 67.2% mean IU on 2012), NYUDv2, SIFT Flow, and PASCAL-Context, while inference takes one tenth of a second for a typical image.

Comments: to appear in PAMI (accepted May, 2016); journal edition of arXiv:1411.4038

Subjects: Computer Vision and Pattern Recognition (cs.CV)

Cite as: arXiv:1605.06211 [cs.CV]