

Human Computation: Object Recognition for Mobile Games Based on Single Player

Mohamed Sakr

Department of Computer Science, Faculty of Computers and Information, Menoufia University, Egypt
Email: mssakr@gmail.com

Hany Mahgoub

Department of Computer Science, Faculty of Computers and Information, Menoufia University, Egypt
Email: h_mahgoub@yahoo.com

Arabi Keshk

Department of Computer Science, Faculty of Computers and Information, Menoufia University, Egypt
Email: arabikeshk@yahoo.com

Abstract—Smart phones and its applications gain a lot of popularity nowadays. Many people depend on them to finish their tasks banking, social networking, fun and a lot other things. Games with a purpose (GWAP) and microtask crowdsourcing are considered two techniques of the human-computation. GWAPs depend on humans to accomplish their tasks. Porting GWAPs to smart phones will be great in increasing the number of humans in it. One of the systems of human-computation is ESP Game. ESP Game is a type of games with a purpose. ESP game will be good candidate to be ported to smart phones. This paper presents a new mobile game called MemoryLabel. It is a single player mobile game. It helps in labeling images and gives description for them. In addition, the game gives description for objects in the image not the whole image. We deploy our algorithm at the University of Menoufia for evaluation. In addition, the game is published on Google play market for android applications. In this trial, we first focused on measuring the total number of labels generated by our game and also the number of objects that have been labeled. The results reveal that the proposed game has promising results in describing images and objects.

Index Terms—ESP game, Games with a purpose, Human computation, object recognition, Smart phones.

I. INTRODUCTION

Nowadays smart phones and its applications gain a lot of popularity. They are accepted and admired by many because they are capable of providing services such as banking, social networking, and fun etc all on the go. Image retrieval (IR) systems contain a database with a large number of images and their labels [1]. The system retrieves the image from the database according to a query entered by the user. One factor that specifies the quality of the IR systems is the images and its labels. Several techniques are used to increase the number of

labels to the images. One of these techniques is games with a purpose (GWAP).

Games with a purpose (GWAPs) are an interesting research field, first presented by Luis von Ahn (2008). GWAPs are one of the human computation techniques, which transform problem solving into game rules [2, 3]. They are used for solving difficult problems using human intelligence. Von Ahn and Dabbish [4] classified GWAP into three game-structure templates that generalize successful instances of human computation games: output-agreement games, inversion-problem games, and input-agreement games. Yuen et al. [5] added output-optimization game to these three templates.

One of GWAPs is ESP game. The ESP game is a two player game that is used to label images. In the ESP game the players are chosen randomly and are assigned the same image. Each player doesn't know the other player and the two players can't communicate with each other. The only thing they have in common is the image that they play with. Each player is asked to give description to that image and has to guess what the other player is typing for each image to win the game and go to the next image. Once the two players have entered the same word, this word becomes the label for the image. The easiest way for both players to type the same string is by typing something related to the common image. The round lasts for 2.5 minutes. During the round the players try to describe as many images as they can. The players get number of points for each image they label. If the players agree on 15 images they get a large number of bonuses in points. Once there is a difficult image that the players can't agree on they both can press the Pass button. The game is attached with a scoreboard, with the names of players with the highest scores. Empirical studies of other peer-production systems have shown that points are a key feature in motivating users [6].

As shown above in the ESP game the more players play the game the more images will be labeled. This will lead to increase the quality of image retrieval systems. To increase the number of players in the ESP, smart phones

and its applications are used. So the ESP game needs to be ported to smart phones this will lead to benefit from the popularity of them. But in the other hand ESP has to support the characteristics of smart phones such as limited hardware resources and portability anywhere any time. One of the characteristics of smart phones is portability anywhere any time. The characteristics of the smart phones have advantages and disadvantages to the ESP. The advantage, as many players will play the game any time anywhere. The disadvantage, ESP game is two or more players which lead to use network to connect players with each other through (Wi-Fi or GSM). The usage of Wi-Fi or GSM will lead to consume the battery of the smart phones as we said before the smart phone resources are limited.

To solve the problem that ESP need two or more players to label images. Jakub developed a new concentration game called PexAce [7]. The game solves the artifact validation problem using “helper artifacts”. The game is card matching game but with modification. Jakub modifies the game by adding a new option for the players to use. This new option allows players to help themselves by writing notes about the images they match. These notes help the players to match the given cards. During the game the player flip each two cards to front side if the two cards match each other they will remain in front side and score will be increased for the player if they don't match each other the two cards will go back to its back side. The player will go on doing this until all the cards are matched with each other. The score for the game is calculated by the total number of flips the player makes. The player needs fewer card flips to finish the game and get more points. The player is allowed to add a description to each card he/she flips. This description will help the player to correctly match the cards with the fewest possible moves. The player had to write the description correctly for these cards to help him successfully match the cards. By modifying the old matching game Jakub are able to convert the game to one player game. Also they solve an important problem called cold start. This problem is due to that upon the initial game deployment number of players will not be sufficient to match each other effectively (especially if the game rule is that the players do not know each other like ESP Game). Some other systems solve this problem by using already annotated resources or known facts [8].

This paper will focus on modifying Jakub method and benefit from it through two ways, first in porting the old matching game to work on smart devices, and second in using it to recognize some objects in the image.

The paper is organized as follows: an overview of related approaches in GWAPs and object recognition is provided in section II. In section III our game and object recognition approach is described. Results and simulation analysis of the proposed approach is presented in section IV. Section V provides conclusion and future work.

II. RELATED WORK

Previously researchers focus on solving the problems that needs a huge processing power and resources using computers. The field of human computation tries to solve the problems that are difficult to computers to solve and trivial for human to solve using human brain power. This field needs human brain to solve these problems. In the past there were two main ways to use human brain:

- 1) Pay for humans to solve the problems but this will be costly.
- 2) Make use of the human desire to be entertained. Convert the problems to a computer game. The humans play the game and in the same time solve the problem as a side effect. These games are known as games with a purpose (GWAP).

In the field of GWAP researchers use computers to solve that types of problems. To gather the maximum number of players they publish the games online on the web to enable any one to play the game from the computer. Nowadays Smart phones and its applications acquires a lot of time in person's life than computers. As a result the previously developed GWAPs need to be ported to smart phones. One of the successful applications of the games with a purpose that harvest human intelligence and time to solve tasks is the ESP game. ESP game is computer game that is used to label images. The ESP game needs to be ported to smart phones and also needs to satisfy the rules of the smart phones. One rule of the smart phones is that the applications on it need not to be resource greedy as the smart phones resources are limited. As a result the applications on the smart phones must not use the resources heavily. The ESP game is a multiplayer game it means that it uses network a lot in its process cycles which will be against the rules of smart phones. The ESP game uses network a lot as it is a multiplayer game, what if it is modified to be one player game?

The authors in [7] solve the problem that the GWAPs need to be multiplayer game by developing a game called PexAce. PexAce was developed to be a single player game to solve a problem called cold start. This problem is due to the lack of players in the initial game publishing stage. As PexAce is a one player game it will be a good candidate to be ported to smart phones. Also we need to add a feature in the game that allows the player to describe not the overall image but parts from it. As a result the player will focus on describing objects that is located in that part only. Peekaboom is one of Ahn's games, focuses on identification of the exact contours of features in images and objects in it [8]. In the GWAPs there has been a lot of work was proposed and a lot of systems were developed. In the next paragraph we try to present some of the previous work that is proposed.

Several other games other than the original ESP game

were devised for a similar purpose [9]. The Kiss Kiss Ban game is a modification of the ESP Game which introduces a third player as an opponent to the remaining two (partnered) players, either to increase specificity of tags retrieved from the game and also due to anti-cheat purposes [10]. In the semantic web domain the researchers in [4, 11, 12, 13] are developed an GWAPs systems to enrich knowledge bases such as ontologies or taxonomies by collecting fragments of knowledge supplied by players. Also a lot researches were developed to help in metadata acquisition which creates metadata beyond the capabilities of automated approaches for textual pronoun co-reference finding [14, 15], images [4, 9, 10] or even people [16].

Several works have been made in the field of image retrieval and image processing. The authors in [17] solve the problem of fast detection for motion using object recognition. For the image retrieval the authors in [18] used the Protégé Tool for ontology development and query retrieval.

III. MEMORYLABEL GAME DESIGN OVERVIEW

Child match, whose purpose is to recognize objects for input images, is a mobile based game that modifies the ordinary concentration card matching game. Concentration card matching game is a turn-based board game for two or more players, who compete to collect the most of the image card pairs, which are mixed and placed on a table facing down (a standard game comprises a board of 8X8 cards). Each turn a player may flip two arbitrary cards to see whether they form a pair. If yes, the player keeps the pair (and receives points for it), otherwise the cards are flipped back and the next player continues. The key to success is to memorize the positions of the cards that have been flipped during unsuccessful attempts (one's own or those of other players) and retrieve them during one's turn. Child match game modified concentration card matching game in three points:

- 1) The game is mobile based.
- 2) Allow players to cheat by adding notes to cards.
- 3) Recognize objects in the input images as a side effect of playing it.

The first modification is that the game is a mobile based game. It runs on the smart phones that support android operating system. The second modification is that the game allows players to cheat by adding notes to cards which make the game useful in many ways. In contrast to concentration, we allow players to cheat by adding textual annotations to images acting as helping notes. Whenever a player flips a card, he is allowed to write a short text on it. This annotation is visible for the rest of the game without the need of flipping the card again. This way, the player can take advantage of his annotation when he encounters the second card of a pair for finding the first one. Therefore, he needs fewer card flips to finish the game and receives more points. After the game, the annotations the player has created during the game describe the input images correctly with high probability. The key to the

descriptiveness of the created annotations lies in the fact, that only a correct description of the image truly helps the player to determine, whether a particular card contains the image he currently seeks. The option of using textual annotations motivates the player to annotate images to improve his score. The third modification is that the input images to the game are not complete images. Each image before entered to the game is partitioned to five parts as described in Fig.1. The images are partitioned in this way as the fact that the main objects of any image lies in the center part of it and the others lies in the corners of the image. This modification allows players to enter descriptions to only parts of the image not the whole image. This will help in the IR systems as when the user searches for a face for example, the IR system will find every image that contains the face tag not only this but also it will draw a rectangle on the part of image that is associated with the face tag.

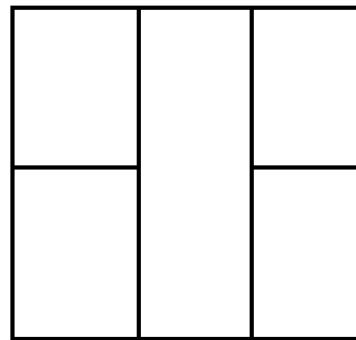


Fig 1. Image partitions

The scoring function has been therefore modified and is based on:

- The number of card flips needed to find all pairs, where fewer flips equal more points.
- The number of tags added to each card, where fewer tags equal more score.

The game is divided into two parts: game on smart phone and on the server as shown in Fig. 2. The first part is the game on the smart phone and its scenario is described as follows:

- 1- The game starts by screen to allow the user to enter his email
- 2- After the user enters his email, this email is sent to the server
- 3- The server finds all the suitable images for this user, and then it sends the images to the game on mobile.
- 4- The game loops through every image that was received from the server and partition every image to five parts as described previously.
- 5- After the partitioning stage the game starts to initialize the game board with hidden cards facing down and score is initialize to 100 (as minimum number of card flips).

- 6- The player is allowed to flip any two cards. When the two cards are matched the two cards are remained visible. If the two cards are not matched they will be flipped back again.
- 7- In any stage the player is allowed to add description to any card by holding in the card. When the player holds on a card a dialog is shown. The dialog contains two images, one is the image associated to that card and the other is the full image that contains that part of image (the whole image and the part). The dialog also contains a list of all descriptions associated to the card and a field that allows the player to enter new descriptions to that card.
- 8- The game ends once all card pairs are discovered. The player is shown the final score. After that the game sends all the descriptions that the player entered to the server with the player score.

The second part is the game on the server and its scenario is described as follows:

- 1- The server receives email for a given user from the game.
- 2- First it selects the images that don't have enough tags associated with them. Then the sever searches to the images that the player has not played with (i.e. didn't described).
- 3- The server finally sends the images with its ids to the game.
- 4- After the game is finished it sends the descriptions the player entered and the player score to the server.
- 5- The server receives this information and adds them to the database.

A filtering stage is applied to the data entered in the database. This stage loops through the database records and first detect if the player cheats. Also it sends all the descriptions to Google translator to check spelling and correct it. The cheating stage is done by checking if for the duration that the player takes to finish the game. And give decision using two parameters time and score. This decision decides how accurate the data is depending on the previous parameters.

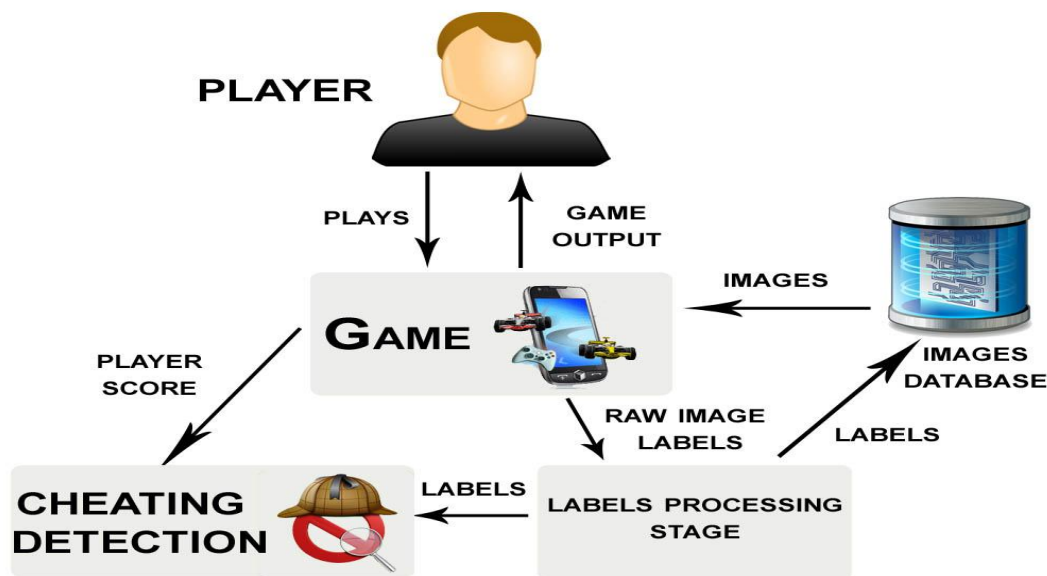


Fig 2. MemoryLabel game block diagram

- 1- Player enters his email
- 2- Player chooses number of cards he wants to play with.
- 3- The game sends the email and number of cards the player choose to the server.
- 4- The server decides the images that the player will play with and sends them back to the player.
- 5- The game starts and a view appears contains cards flipped on back face and current score of player.
- 6- The player is allowed to flip any card to see the image behind it and he has to match it with another card.
- 7- In any stage the player is allowed to enter description to any card by holding down on it without flipping it.
- 8- After the player finish matching all the cards the game sends the score of player and all the labels he enters for the cards to the server.
- 9- The server handles the labels and player score to decide if the player cheats.
- 10- Also the server sends the labels to Google translator to check spelling.
- 11- At the end the server inserts the labels and player score into the database.

Fig 3. MemoryLabel game work steps

An IR system was developed to test the data entered in the database. The user enters a search query then the system looks inside the database and retrieve every image associated with the entered query. The system also

retrieves every image with the tags and the part number of the image that the tag associated with it. Also it retrieves the score of the player that enter the description with the highest possible score for every tag where the score is the

number of moves the player make to match two cards. The system displays the resulted images to the user with a rectangle above the part of image with the associated tags. The work steps of the MemoryLabel game can be summarized as shown in Fig.3.

IV. RESULTS AND SIMULATION ANALYSIS

At the beginning we start to get a data-set of images and labels to work on. A number of crowdsourcing data-sets are available for research. For example, von Ahn et al. contributed a list of 100,000 images with English labels from their ESP Game [19]. We used von Ahn et al. data-set. First, we integrate the images to our system on the server. We deploy our system at the University of Menoufia for evaluation using Java 1.6 as programming language and android to develop the mobile game memoryLabel, Mysql [20] server as database management server, php programming language for developing the web service and run the server over Intel core i7 with 4 GB Ram PC on windows 7-64 operating system.

The trial of memoryLabel is shown in Fig.4. In this trial, we first focused on measuring the total number of labels generated by our game memoryLabel. We sent emails to the staff of the University. We advertised the system as a free game and provided a link to the game. Five users signed up initially (many of them were research students and demonstrators). The five players played on a 13 images. Some of the players play on one image in the game round and some of them played on more than image in the game round. All the five players who play with one or more image in the game round added 48 labels for describing the 13 images. Some of the labels were very informative like hand, hair, eye and so on. The other labels were describing only colors like black. The total 48 labels were verified by the author manually without any helping software. Some of them had spelling errors and other was correct. The results of the memoryLabel game are listed in Table I.

TABLE I: THE RESULTS OF THE MEMORYLABEL GAME

Total players	Total Images	Total Labels
5	13	48

Due to the time limitation, we did not observe the users for longer period to give more answers. However, we believe if we deploy the system to a larger demographic, our game scenario would produces even more promising results fuelled by the network effect.

A. Case study of the memoryLabel game results

This case study illustrates the results of the memoryLabel game as shown in Fig.5. The user is presented with an image which is partitioned into 5 parts; the five rectangles in red; as shown in Fig. 5. The player give description to each part for example the upper left part the player give the words (wall, hair) as description

for that part and so on. This enable the IR system to give the user information about not only the image contain the word the user enters but also where that word is located in the image.

V. CONCLUSION

This paper presents a new game called memoryLabel to give descriptions to objects located in the images. The game uses the advantages of smart phones popularity. The game modifies PexAce game to be a mobile based game and also to describe parts of the images not the whole image. The results of the memoryLabel game reveal that the game can help in recognizing objects in the images and give a description of these objects.

In the future work we intend to divide the image not into five parts but depending on edge detection and object extraction algorithms from computer vision. Furthermore we intend to evaluate the performance of our game in terms of user experience to improve it.

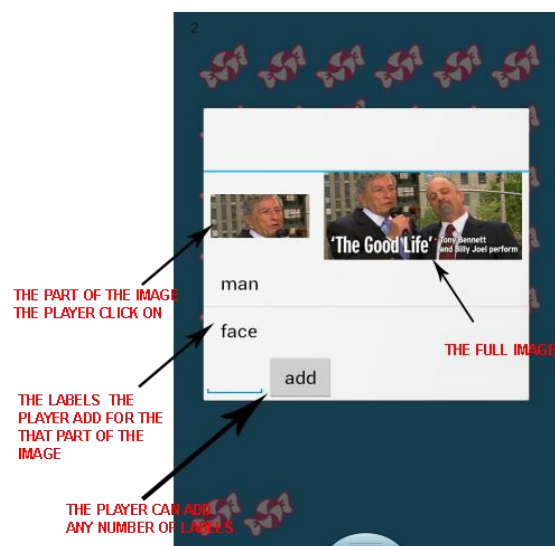
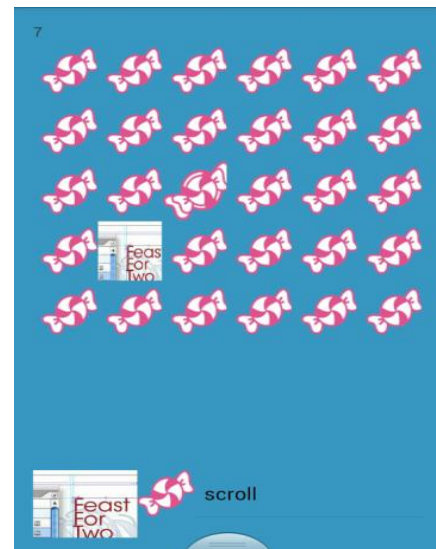


Fig 4. Trial of the MemoryLabel game

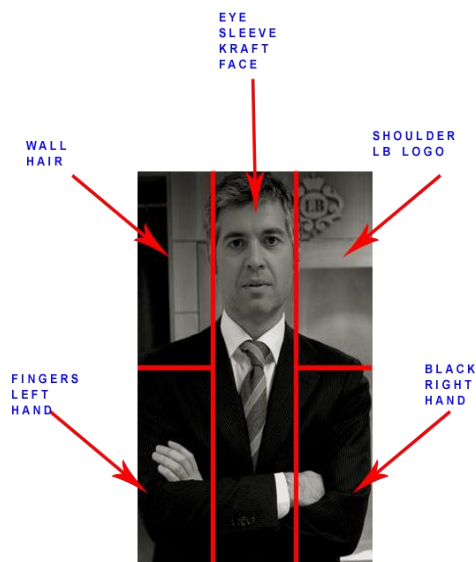


Fig 5. MemoryLabel game case study

REFERENCES

- [1] Datta R, Joshi D, Li J, Wang J Z. Image retrieval: Ideas, influences, and trends of the new age. *ACM Computing Surveys (CSUR)*, 2008, 40(2).
- [2] C. Harris and P. Srinivasan, Comparing Crowd-Based, Game-Based, and Machine-Based Approaches in Initial Query and Query Refinement Tasks, *Lecture Notes in Computer Science*, 2013, 7814: 495-506.
- [3] S. Thaler, E. Simperl and S. Wolger, An Experiment in Comparing Human-Computation Techniques, *Internet Computing*, IEEE, 2012, 16(5):52-58.
- [4] Ahn, L.v., Dabbish, L. Designing games with a purpose. *Communications of the ACM*, 2008 51: 58–67.
- [5] M. Yuen, L. Chen and I. King, A Survey of Human Computation Systems, *Proceeding CSE '09 Proceedings of the 2009 International Conference on Computational Science and Engineering*, 2009, 4: 723-728.
- [6] K. Nam, M. Ackerman and L. Adamic, Questions in, Knowledge in ? A Study of Naver's Question Answering Community, *Conference on Human Factors in Computing Systems, CHI 09*, ACM, 2009, 779-788.
- [7] Simko J, Tvarok M, Bielikova M. Human computation: Image metadata acquisition based on a single-player annotation game. *International Journal of Human-Computer Studies*, 2013, 71(10): 933-945.
- [8] Ahn L V, Liu R, Blum M. Peekaboom: a game for locating objects in images. *CHI '06 Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 2006. 55-64.
- [9] Seneviratne, L., Izquierdo, E. An interactive framework for image annotation through gaming. In: *Proceedings of the International Conference on Multi-media Information Retrieval*. ACM, New York, NY, USA, 2010, 517–526.
- [10] Ho, C.J., Chang, T.H., Lee, J.C., Hsu, J.Y., Chen, K.T. KissKissBan: A competitive human computation game for image annotation. In: *Proceedings of the ACM SIGKDD Workshop on Human Computation—HCOMP '09*. ACM, New York, NY, USA, 2009, 11–14.
- [11] Krause, M., Takhtamysheva, A., Wittstock, M., Malaka, R. Frontiers of a paradigm - exploring human computation with digital games. In: *Proceedings of the ACM SIGKDD Workshop on Human Computation—HCOMP '10*. ACM, New York, NY, USA, 2010, 22–25.
- [12] Markotschi, T., Völker, J. GuessWhat?!—Human intelligence for mining Linked Data. In: *Proceedings of the Workshop on Knowledge Injection into and Extraction from Linked Data (KIELD) at the International Conference on Knowledge Engineering and Knowledge Management (EKAW)*, RWTH, 2010, 28–39.
- [13] Siorpaes, K., Hepp, M.. Games with a purpose for the semantic web. *IEEE Intelligent Systems*, 2008, 3 (23): 50–60.
- [14] Hladka, B., Mirovsky, J., Schlesinger, P. Designing a language game for collecting coreference annotation. In: *Proceedings of the Third Linguistic Annotation Workshop*, Association for Computational Linguistics, Stroudsburg, PA, USA, 2009, 52–55.
- [15] Chamberlain, J., Poesio, M., Kruschwitz, U. A demonstration of human computation using the Phrase Detectives annotation game. In: *Proceedings of the ACM SIGKDD Workshop on Human Computation—HCOMP '09*. ACM, New York, NY, USA, 2009, 23–24.
- [16] Guy, I., Perer, A., Daniel, T., Greenshpan, O., Turbahn, I. Guess who?: enriching the social graph through a crowdsourcing game. In: *Proceedings of the 2011 Annual Conference on Human Factors in Computing Systems—CHI '11*. ACM Press, New York, New York, USA, 2011, 1373–1382.
- [17] Khalid Y, Peng-Yung W. Difference of the Absolute Differences – A New Method for Motion Detection. *International Journal of Intelligent Systems and Applications (IJISA)*, August 2012, 4 (9).
- [18] Vishal J, Mayank S. Ontology Development and Query Retrieval using Protégé Tool. *International Journal of Intelligent Systems and Applications (IJISA)*, August 2013, 5 (9).
- [19] ESP Game dataset: <http://server251.theory.cs.cmu.edu/ESPGame100k.tar.gz> , accessed Oct 2013.
- [20] MySQL Server <http://dev.mysql.com/downloads/mysql/> , accessed Oct 2013.

Mohamed S. Sakr received his B.Sc degrees in Computer Science from Menoufia University, Egypt in 2011. His research interests lie in the areas of human computation, data mining and cloud computing and currently he is demonstrator at Computer Science department in the Faculty of computers and information, at the University of Menoufia, Egypt.



Hany Mahgoub received the B.Sc. in Pure Mathematics and Computer Science and M.Sc. in Computer Science from Menoufia University, Faculty of science in 1993 and 1999, respectively. He received his PhD in Computer Science from Menoufia University, Faculty of Computers and Information, Egypt in 2010. His research interest includes data, text and web mining, Information Retrieval, software testing, distributed systems, intelligent systems, human computation and bioinformatics.



Arabi E. keshk received the B.Sc. in Electronic Engineering and M.Sc. in Computer Science and Engineering from Menoufia University, Faculty of Electronic Engineering in 1987 and 1995, respectively and received his PhD in Electronic Engineering from Osaka University, Japan in 2001. His research interest includes software testing,

software engineering, distributed system, database, data mining, and bioinformatics.