GWAP Final Report

Piyush Rumao
Xuanhui Xu
Yuhang Liu
COMP47580

Introduction

The main purpose of GWAP (Games with a purpose) is to build a computer game that is fun to play but contributes to solving real-life problems at the same time. And this is normally realized by carrying out the idea of collective intelligence, which aims to involve a great amount of participants to reasonably increase the accuracy of the final outputs. GWAP is a perfect example of applying collective intelligence --- every single player of the game contributes to solving the target problem. Besides, GWAP is perfect for solving the problems which are extremely difficult for computers but easy for humans, .e.g image recognition. A good example of a GWAP is the ESP game. In the ESP game, players are randomly paired at the beginning of the game, and they are asked to label the image shown on the game panel. If both players have labelled the image with the same word, or in other words, used the same word to describe the image, they win the round. The ESP game uses the images from the internet, and as players play these images are labelled with various words. And because the label is from two independent sources, so they are usually very good labels for the image.

Inspired by the ESP game, this project is a multiplayer sketch-labelling game with the purpose of enhancing image recognition technologies. Players are expected to sketch a certain object (or a scene) in turn, and the rest of the players are supposed to guess what the drawing is. The gameplay is supposed to be asymmetrical, as only the player who draws gets to know that the topic is. Having managed to guess the answer or having others managed to guess the answer gives you scores, and the one with highest scores wins. Furthermore, there would be three levels containing in the application and user needs to unlock other levels by finishing the previous level and getting enough points. In the first mode, player can only use points to describe the topic given, while lines and faces are forbidden. Similarly, the second mode does not use points and lines but only color blocks. Finally, the mode is to distinguish whether the describer is human or not. In order to active players enthusiasm, a leader broad will be set based on the total points player gets from every mode.

In this report the motivations will be discussed, as well as the target problems that we're trying to tackle. And the main body of the report will be the detailed game designs and mechanics, illustrating how the game is played. In the end the report the project will be the evaluation including some additional detail on how the efficacy of the project is evaluated, following by the conclusion.

Motivations & Target Problem

"When you draw an object, the mind becomes deeply, intensely attentive," says the designer Milton Glaser, an author of a 2008 monograph titled *Drawing Is Thinking*. "And it's that act of attention that allows you to really grasp something, to become fully conscious of it." With increase in chat application development for smartphones and ease to access internet which makes it even more easier for people from every corner of world to connect with each other, developing a application like one seems very reasonable which makes the game playing easier, pulling out more crowd as it is fun to play and added advantage that people get in practicing doodling.

Now, comes the Backend Purpose of the Game which already has so many front end advantages. Doodling is coming out to be new upcoming trend where people can easily sketch a million dollar idea that occur in their mind on their tablet. In such cases our current dataset are not strong enough to detect collaborative meaning of the painting, so we need a more valuable dataset that holds labels of individual doodles as well as their meaning if associated as a whole.

The third motivation to go for doodling was that current image processing algorithms are very specific and give an incorrect result even if there happens to be a slight distortion or noise in the image, while the high end image is been downgraded in lowest form for processing. But if we train our system to identify an object based on multiple sketches of the object our system trained will be much more robust and thus can easily tolerate little distortion in image to be processed. Doodling on a canvas, unlike images, have a potential of going "wild". Even the most common object may be drawn in most divergent ways. And because the doodling set is generated by human players, the algorithm could grow to be of great practical significance.

After years of development, image recognition is becoming a more and more mature technology. However, certain types of images, doodling, for example, are not quite recognizable for AIs. When we draw things, we naturally extract the most prominent features of the object. For example, if a person is asked to quickly draw a mallet, he or she will most probably draw two rectangles adjacent to each other. And shape combinations like this normally makes sense to us human, but relatively not that much to artificial intelligence algorithms. Extend this idea a bit further, objects drawn in colored blocks or in a dot connecting style can also do the trick. Algorithms are able to learn way faster than humans, but when it comes to tasks like this, human brains are way more effective. Hence with the help of the game players, our algorithm could learn to recognize doodles over time.

Since AI can learn from human sketches, could it be as good as us one day? Here comes the interesting part of our game. In our game, AI may also play the role of describer. Therefore, when the drawer sees the response from his opponent, he'll have to distinguish if it comes from a human player or the algorithm.

In order to make artificial intelligence have the ability to recognize human sketching and doodling, plenty of labeled images are needed for machine training. As neural network can label the basic sketching for common stuff, the requirements of recognition in classification and subjects has risen for instance, to identify how close are machine learning algorithm capable of labelling a scenario when given as a whole (collaborative meaning hidden in picture) and do they follow the human emotional values and viewpoint that can be associated with a collective picture. Furthermore, how well could artificial intelligence

perform is also needed to be evaluated or Are humans effective enough to identify the difference between a bot and human being if jumbled up in an environment where they are unaware of the opponent.

The main challenging points of this target problem are basically two aspects: One is the origin labelling work can only be finished by human. The other one is only human can judge whether a bot is a bot or not. By using GWAP-approach, we can easily achieve all the motivations (doodling database, labelling and performance evaluation) into one type of game and let lots of players to provide their own thoughts as our data.

The system will collect the doodling images from every game with the correct label. The data collected is image and words, so they are definitely usable, which could be used in other stages of the game. On the other hand, the argument to artificial intelligence from player can tell whether it should enhance or decline the current judgement.

Related Work

The question of object–picture recognition has received relatively little attention in both human and comparative psychology; a paradoxical situation given the important use of visual technology; made by neuroscientists in their experimental investigation of visual cognition. Taking this into consideration Dalila Bovet , Jacques Vauclair in their research, “Picture recognition in animals and humans” conducted experiments proving that the nature of a picture (e.g. bi-dimensionality, reduction of cues related to depth), it is suggested that object–picture recognition be envisioned in various levels, and thus giving us a go ahead to state that image doodle can prove to be much better way to train a system to recognize a situation and at same time help humans improve memory thus giving a win-win situation for our GWAP application.

Joaquim A. Jorge, in his paper, “A Simple Approach to Recognise Geometric Shapes Interactively”, researches about how humans recognize complex structure and shapes even by recognizing the line geometry and what possible algorithm can be applied in such case if we wish to get our artificial neural learning algorithm to be as good as humans in recognizing the shape as whole even if given as canvas of incomplete lines. Keeping this in mind we created task three which gave opponent an incomplete canvas of disjoint lines and user has to join those lines to create a picture as a whole.

Joseph Redmon, Santosh Divvala, Ross Girshick, Ali Farhadi in their most famous research, “Real-Time Object Detection”, talks about object detection and then labelling them into particular objects using artificial neural network which has been widely implemented into google brains team, which can be used to identify individual objects, where they have billions of dataset of individual object images been hand labelled which are been used to train the model and then identify object distinctly, taking this research ahead we plan to create dataset to label the canvas as whole which will derive the collaborative meaning of the whole painting. So we can use these label images to identify not just individual objects but as whole meaning associated with it. Thus creating a whole new repository for it using these GWAP and opening new doors for research in these field.

Sue Shellenbarger, in her article in “The Wall Street” talks about Recent research in neuroscience, psychology and design shows that doodling can help people stay focused, grasp new concepts and retain information. It even plays a key role in stress relieving with more of such articles coming out in the Wall Street and similar such magazines, getting an application in such a field will definitely see a sudden increase in users just as we talk of our business plan which means gathering maximum users who will play the game and thus we will be able to generate a sufficiently large dataset.

Jamie Condliffe, in his article on Google AutoDraw, which talks about doodle processing where AutoDraw can convert your doodle into a meaningful and proper Doodle, these idea played a key role in deciding sketches provided by naive users and professional users and compare results of both on same query so that we can further use similar technology to train system fill in empty spaces and create a correct sketch as a whole.

Game Design & Mechanics

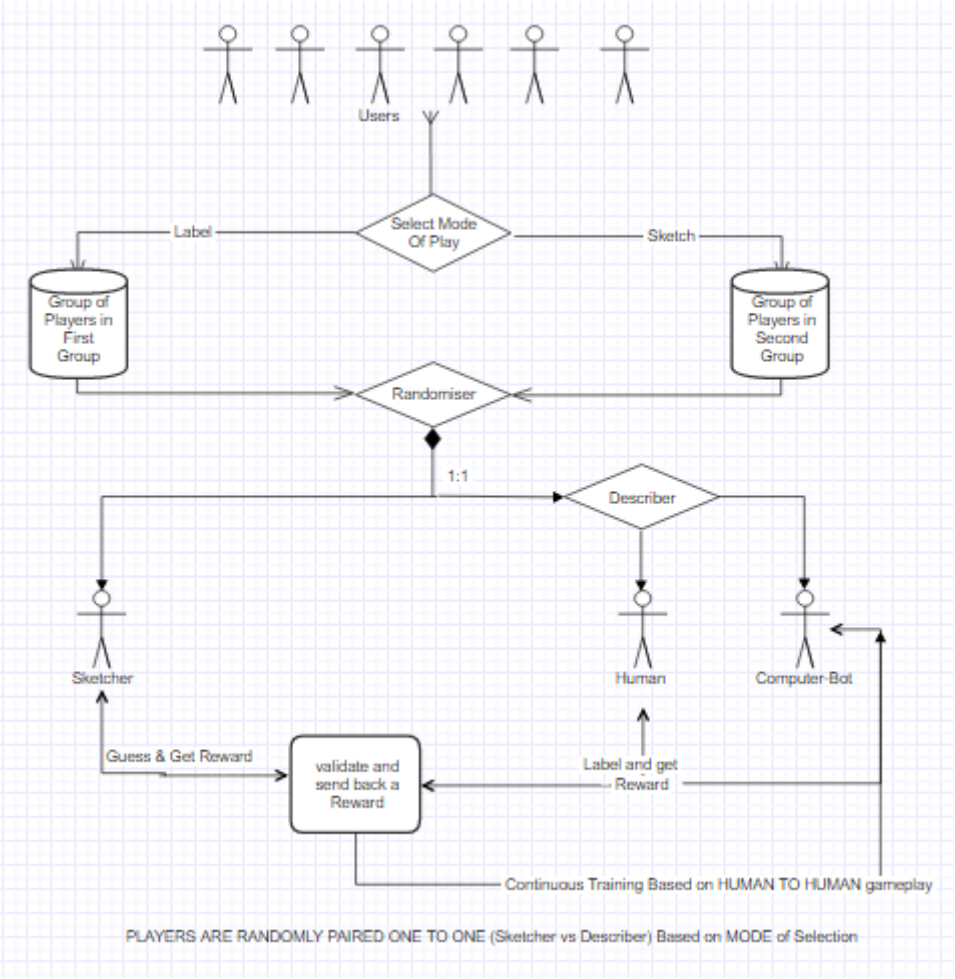
It's a multiplayer sketch-labeling game. Players are expected to sketch a certain object (or a scene) in turn, and
the rest of the players are supposed to guess what the drawing is. Or, in the player-against-AI mode, players will have to guess whether it's a human or AI whom he's playing with. The gameplay is supposed to be asymmetrical, which means in all of the game modes only the drawer knows the topic. Having managed to guess the answer or having others managed to guess the answer gives you scores, and each player has a personal all-time score which will be used to compete against each other on the leaderboard.

The game starts with a pairing process, during which players are randomly paired with each other. Inside the game, it contains three modes with different gameplays. The first two modes is simply draw & guess. One player will be given a topic, which could be either an object or a scene, and others have to the correct answer. And the faster you make the correct guessing, the higher scores you will get for this round and as for the player who draws, the number of correct guesses contributes to his/her scores. Thus everyone who doodles will have to try to draw well if they want to have higher scores. For each set of game there will be many rounds, and the one who scores the most wins.

The doodling comes in different types. In addition to basic, simple sketching, the doodles can also be presented in shadings. In this case players will have to guess what the drawing is by its color instead of rough sketch. What's more, the previous drawing could be reused. These sketching could be presented to players in different ways, for example, the sketching could be presented separately, a random part at a time. Or it could be presented as a whole, but in a blurred effect. The players are still required to guess the drawing, but in this way
more information could be generated to help with our purpose. Another mutation of doodles is dot outlines. Just like the dot connecting games, the object's outline is drawn in dot lines, which makes the game even more challenging. The first two modes only support doodles in color blocks and dot outlines, and the third mode supports regular doodling.

What could bring the game to another level, and exactly what makes our game unique is the third mode, which is a simplified version of Turing test. An artificial intelligence will join the game and try to guess what the drawing is, just like a regular human player. And the AI could either be describing or drawing. This will add much more fun to the game, because it could be more and more challenging as the AI "learns" through time. What's more, the user inputs also train the algorithm to make it more human-like.

Let us have a look at the architecture diagram of our Game below:



So, our Game begins with multiple users who can select the mode that they wish to play the game in, by selecting either they wish to sketch or label a sketch within a given time-frame. Based on this selection our system categorizes the current users into two active groups, which are passed on to a randomizer module which associates two users, each from different group into a single game-mode or can even add a human vs bot competition. Once the initialization phase is completed here, we have a sketcher and a Describer (Human or Bot) playing against each other. Now our sketcher gets a task to draw, and he has to complete it in a given time, for example: say Draw an apple in 30 seconds (can be more complex too). Now the describer within this time frame has to keep labelling the image before the sketcher is finished and based on his labelling the sketcher validates his labels which in turn will send reward points to the describer, Now sketcher while grading the describer has a additional section to guess who the describer was, was it a human or a computer Bot, based on his answer the sketcher gets a reward if he guesses it correctly. We use this step of Sketcher and Describer (Human vs Human) to train our bot to be even more efficient with every possible game-play which will make it difficult for sketcher to distinguish between human or bot. Now to make the game interesting we have three modes. We are attaching the screenshots below.

As there are three modes in this application, three buttons respectively corresponds to Dot mode, Color Block mode and AI mode. At the beginning of the game, player can only enter the Dot mode. Next, when player's total score reaches 50 points, the Color Block mode will be unlocked. Similarly, the last mode needs 100 points to unlock. The big score and leaderboard are designed to act user's motivation and enthusiasm. With higher points, player would be given more stars to show off their level of this game. Furthermore, player is able to change the volume and sound effect in the setting page. If they want to hide from the leaderboard, they can also edit in setting.

The first mode is Dot Out Outline. In this mode the doodle is given in a dot-connecting game style as shown in the figure. In the example above, the apple object is outlined in dot lines. Please notice that in this mode it's still possible to draw colored dots, otherwise it might be too difficult to guess. Other players can guess the object by typing it out, however, they only got one chance. Because of this, in some complex topic, waiting for other players guess first might be a good strategy to win the points. All player's answer would be showed on the screen for referencing. In the end, the score player wins or loses will be summarized on one board. If none player can figure out the doodling within a period of time, the sketcher may not get any points in this round.

The second mode is Color Blocking mode. In this mode the radius of the painting brush is set very thick, making it almost impossible to draw the details of the object. Hence the objects need to be represented by chunks with colored. In the example above, the apple object is drawn with a big red circle, an orange chunk and a smaller green circle. The other rules are nearly the same as the previous mode, but it is mainly challenge player's ability of abstracting object by color blocks. By using the thick lines to express what player wants to express, the difficulty of the game has been pushed to another level. Participants need to figure out a way to win more points in this mode as the value of this mode is higher, or they may not get the ability for the last mode.

Finally, in the third mode "Who Replies You", players will be playing against another human player or a bot player. In this mode the doodles are delivered in a more "general" way, otherwise the AI is not able to recognize the contents. When the drawer finishes doodling and has the answer from his opponent, he's supposed to make a guess on the identity of his opponent -- whether it's a man or a bot. At the very beginning, the player is asked to choose to be a sketcher or a describer. As a sketcher, player need to doodle the picture and guess whether the opponent is a human or not. Only if they get the correct answer, they can get points in the end. While as a describer, they only thing they need to do is to tell the sketcher right answer or what human is thinking about this doodling. The trick way of showing the identity is forbidden in this mode, such as typing "I am human" to tell the sketcher. In order to avoid this happening, the AI would randomly type those obvious trick words to confuse the sketcher.

Evaluation:

As it will be a multiplayer game played at different level of complexity, we need to implement different evaluation measures.

To begin with, we need to first examine the overall design of the game: is it fun to play? are the instructions all clear and comprehensible? can we achieve the purposes of the game? These are some of the very crucial factors that we need to take into consideration. The first two questions should be evaluated by introducing the game to a vast number of users and collect the user feedbacks. Gaming experience is a rather subjective measurement and cannot be simply quantified. Ideally the players should be able to figure out how to play at their first attempt. In addition, the game should be competitive but fair to players with different backgrounds.

For the major functionalities of the project, there are several stages of evaluations that need to take place. Stage 1 will identify how well does our bot label human sketches based on previous training. which will

even be extended to level ahead to say does our Machine Learning algorithm considers the collaborative meaning associated with the picture as whole as humans do, which means getting a sense out of collection of certain types object by training our bot to learn from occurrence of certain pattern of objects together. This can be achieved only when we have large amount of quality data, which can be achieved by diverse amount of people participating in game play.

Stage 2 will evaluate how good are humans at identifying their opponents based on labelling structure if its human or a bot, when made to play multiple games with different opponents. It shouldn't be either too easy nor too difficult to maintain the gaming experience. More importantly, a badly-tuned AI is no good for its self-learning processes.

Stage 3 will evaluate humans perception ability by the way they fill the empty blank to create a complete doodle based on their mental image of what an incomplete doodle (given to them on screen) can be or label a doodle correctly if given upside down. And these evaluations mentioned here all focuses on the accuracy of guessing. The accuracy should fall into an acceptable range, neither too low nor too high.

Stage 4 will comprise of actual reward that our users get on answering correctly or completing the task within time-frame, which will generate a sense of competition among the users to gather maximum reward points, thus improving a general users input level in terms of sketch quality, thus a naive user technically will not contribute that great as a medio-care or professional user. Thus in terms of evaluation of their results we can associate an additional tag with the canvas to train our system later that a naive users sketches will be more difficult to interpret than a professional user thus using anti-learning policy where we train system first on good sketches and then slowly make it learn on sketches which might lack in clarity.

We also have to look at the potential cheating points. For example, in the player-against-player game mode, the player who draws could write the answer on the canvas instead of drawing the actual image. This is a major drawback of using a canvas and let players draw --- the drawing contents cannot be effectively controlled or supervised. What's more, they have every reason to do so to get better rankings on the leaderboard.
Unfortunately, this is almost inevitable because of the openness of the game. A possible solution is to build a "report" mechanism. Players who violates the rules can be reported by his opponent, although this is possibly not as effective as it should be. However, we believe this would not appear very often and would only damage the game on a relatively small scale. After all the majority of game players play for fun. As for the player-against-AI mode, human players could write messages like "I am human" to identify himself. A possible way to tackle this problem is to make our AI algorithm send similar messages every so often. This could interfere with the cheaters, but at a cost of sacrificing the gaming experiences of normal players --- guessing such messages is gambling.

Last but not least, it's also of great significance to examine the learning efficiency of our neural network. And in order to do this, the accuracy of guessing should be well logged for generating graphs showing how the performance of the AI changes over time. This includes both identifying human doodles and drawing sketches. These evaluation methods mentioned are supposed to be carried out for many times

as the program is developed, especially when new features are added. We believe it's important to follow a modern software development approach to build software with good qualities.

Conclusion:

At this stage of developing, our game is fully playable. Before the game starts players will be randomly paired based on the game mode they've chosen. Inside the game there are several game modes players can choose from, including regular doodling, color blocks doodling and dot outline doodling. And each game round could be played between human players or human and bot players. Regardless of which game mode the player is playing with, his personal all-time score is recorded and put into the leaderboard to compete against others. Theoretically this mechanism will encourage players to play more and to give more valuable answers. With the enormous amount of doodling data input by players, our neural network is able to gradually learn to recognize doodles. This we consider is a great step ahead from the most famous MNIST dataset to train neural system learn on human doodle and go ahead to deduce a collaborative meaning associated with those doodles.

In addition to the functionalities we have already implemented, there're still quite some improvements that could be done in the future.

Future work

We currently have one to one game play where at a same time only two participants can be connected together, but in future we can plan to have a multiple players been connected together just like a group playing together and there is a chance of one of them been a bot, which will make the game even more interesting a fun to play with as each player will get an equal opportunity to participate in each task of sketching and labelling at same time finding out the unusual player in the group.

One major improvement on the project could be using phrases and words rather than words only to describe the scenes. In this way the description is closer to natural language and, apparently, much more accurate. This also contributes to AI's natural language process abilities.

Bibliography

1. Jonghyun Choi, Abhishek Sharma, David W. Jacobs and Larry S. Davis: Data Insufficiency in Sketch Versus Photo Face Recognition

2. Hu Han and Anil K. Jain: Tattoo Based Identification: Sketch to Image Matching.

3. Jackie Andradre: What Does Doodling do?

4. Google Image Labeller: https://en.wikipedia.org/wiki/Google_Image_Labeler

5. Karol Gregor,Ivo Danihelka,Alex Graves,Danilo Jimenez Rezende ,Daan Wierstra: A Recurrent Neural Network For Image Generation.

6. David Ha , Douglas Eck: A Neural Representation of Sketch Drawings

7. Joaquim A. Jorge: A Simple Approach to Recognise Geometric Shapes Interactively

8. Dalila Bovet , Jacques Vauclair: Picture recognition in animals and humans

9. Joseph Redmon, Santosh Divvala, Ross Girshick, Ali Farhadi: "Real-Time Object Detection"

10. Sue Shellenbarger: "The Wall Street Journal" "https://www.wsj.com/articles/the-power-of-the-doodle-improve-your-focus-and-memory-1406675744

11. Jamie Condliffe: “Google’s AI to turn your crapy Doodles into proper Pictures”
https://www.technologyreview.com/s/604156/googles-ai-turns-your-crappy-doodles-into-proper-pictures/
12. Sunni Brown: The Doodle Revolution
13. Milton Glaser: Drawing is Thinking