

# A Human Visual System Based Temporal Model for Semantic Levels Categorization

MOHAMMAD HOSSEIN KARIMI<sup>1</sup>, REZA EBRAHIMPOUR<sup>2,3</sup>, AND NASOUR BAGHERI<sup>1,4</sup>

<sup>1</sup>Faculty of Electrical Engineering, Shahid Rajaei Teacher Training University, Tehran 16785-163, Iran

<sup>2</sup>Faculty of Computer Engineering, Shahid Rajaei Teacher Training University, Tehran 16785-163, Iran

<sup>3</sup>School of Cognitive Sciences (SCS), Institute for Research in Fundamental Sciences (IPM), Tehran 19395-5746, Iran

<sup>4</sup>School of Computer Science (SCS), Institute for Research in Fundamental Sciences (IPM), Farmanieh Campus, Tehran 19538-33511, Iran

Corresponding author: Nasour Bagheri (nbagheri@sru.ac.ir)

**ABSTRACT** Humans can categorize an object in three ways. For example, a car can be categorized as vehicles (superordinate), ground vehicles (basic), or cars (subordinate). Different semantic levels of categorization are referred to these three categorization modes. There are different speeds and accuracies for a similar object in the classification of these levels. However, much research has been done in this context, the trend of these levels is still questionable as to the accuracy and reaction time and the reason for this difference. In this paper, we examine the order of these levels and the reason for their differences. We show the superordinate advantage is declared, and after this level, the base level and subordinate level are expressed, respectively. To this end, first, we design an experiment to examine the semantic levels in the human eye system. The result of this experiment is the superordinate advantage. In fact, at the superordinate level at the same time, the reaction time was lower than other levels, and the efficiency was higher than other levels. Besides, a computational model is introduced that has time and can classify semantic levels. The model is trained for ten categories in this study. These ten categories are considered a subordinate level, and five levels of basic and two levels of superordinate are expressed. We found that at higher levels such as superordinate, more neurons participate in the clustering task, so the outcome is faster and more accurate results. Moreover, the proposed model was tested for inverted input images to verify the model.

**INDEX TERMS** Semantic levels, superordinate-level, basic-level, subordinate-level, reaction time, object categorization.

## I. INTRODUCTION

The processing time of visual cortex is a useful tool to study the different layers of visual cortex and to understand the hierarchy of the visual cortex [1]. The processing time of different areas of the visual cortex is depended on the problem types. Besides, image processing is quickly carried out in the visual cortex and can be divided into two parts, the primary processing and the high-level processing [2]–[5].

The visual cortex is feed from the image information when a stimulus is an onset. It is shown that the human visual system can immediately react with an average reaction time from approximately 250ms to 400ms [6], while the processing in the visual cortex occurs in a bottom-to-top procedure [7]. In fact, Information first switched by the lateral geniculate nucleus (LGN), and in terms of information hierarchy in the

visual cortex, V1 is the first area that accesses the information. Then, the information in ventral pathway is feed to V2 and V4. Afterward, it is feed to the Inferior temporal (IT) section containing posterior inferior temporal (PIT) and anterior inferior temporal (AIT) two sections. The IT section which is the last section of the visual cortex, contains several neurons which only spike to some specific stimuli. Finally, information is feed to the prefrontal cortex (PFC) and motor cortex (MC) sections of the brain. The delay in the last two sections is considered constant [8]–[11]. Semantic levels are divided into three levels of superordinate, basic, and subordinate [12]. Superordinate is known as the highest level of category and may keep a high degree of abstract information and generalities. The basic level is the fundamental level of categorization; however, it is not able to offer any particular configuration and gestalt. Most object categorizations in human's daily life are placed at this level. The basic level, which is placed between two other levels, is about the

The associate editor coordinating the review of this manuscript and approving it for publication was Junhua Li<sup>1</sup>.

objects' existence, and most common features are categorized under this level. Finally, the subordinate level can be seen as the lowest categorization level which classifies the objects in the lower grade of generic. These three levels of categories perform in high resolutions and details. For example, one can categorize an object as a vehicle, ground vehicle, or a car (superordinate, basic, and subordinate levels, respectively).

Several psychophysical studies have shown more rapid perceptual access to basic-level category information, called basic-level advantage, than to the higher or the lower levels [13]. Nevertheless, some other psychophysical research results are challenging basic-level advantage by indicating faster access to superordinate level category information [14], [15]. Therefore, exploring the temporal sequence of object categorization at different levels of hierarchy is one of the key points to understand the neural mechanism of object categorization in the brain [16]. At a glance, humans can detect an object, recognize it as an animal or non-animal, categorize it as a marine animal or domestic animal, or finally identify it as a dog or cat. The crucial point is to know whether categorization time affects the level of categorization. On the other hand, if this is true, which category is detected earlier?

In 1976, Rash et al. first raised the question in which the classification levels would occur faster [17]. They conducted an experiment in which several labels such as 'Animal', 'Bird' and 'North Cardinal Bird' were shown to a subject and requested to sort them based on shown label. It is shown that basic level was faster and stronger than other two levels.

Another well-known study [18] shows that the categorization pathway is separated at different levels. In the experiment, participants should categorize the stimuli in upright and inverted states at two levels of superordinate and basic. Surprisingly, the accuracy of results is very high and the same at the superordinate level in both states of upright and inverted; however, the accuracy drops at the basic level at the inverted state. The presence of different paths in the classification is highlighted as a reason for this issue.

Another interesting research on categorization levels explains that categorization in the visual system is in that way which categorization performance is based on perceived similarity relations among items within and outside the category [19]. The results in this study were time-consuming for classifying distractors and atypical objects, and most importantly, they have specified the trend of superordinate, basic, and subordinate due to speed and accuracy. They explained that the reason for this trend was the different speed of course and fine information.

A recent study [14] illustrates a variety of experiments that, in all cases, the superordinate is faster. Examined states include variation in stimuli and stimuli onset time, which in all experiments superordinate are reported faster.

In 2015, a study [20] examined the pace of categorizing levels. Numerous experiments reveal that the classification pace of different levels is associated to the stimulus time. Therefore, it is concluded that by increasing the display time, the superordinate advantage becomes to basic advantage.

Besides, another study [21] examined neural responses of the inferior temporal (IT) cortex related to macaque monkeys seeing a large number of object images. It is proposed that the basic level advantage is compared to other two levels.

A recent study [22] investigates the role of spatial frequency in object categorization at different levels. They separated some images into three filters of low spatial frequency (LSF), intermediate spatial frequency (ISF), and high spatial frequencies (HSF) bands and examined three levels of categorization with an HMAX model. In that research, they were explained for basic and subordinate levels, higher spatial frequency is needed, and also accuracy in the basic level is higher than the subordinate level. Moreover, at the HSF band, all levels yield a similar accuracy. The study introduces the stimulus frequency due to the speed of categorization of levels. A review of the levels of categorization can assist to understand the functionality of human eye system. It can potentially explain the classification mechanisms in the visual cortex and address the questions that which level of classification is faster? And also, which level is more accurate?

In this paper, first, we indicate a psychophysical experiment performing on ten participants. We asked the participants to categorize vehicles, ground vehicles, and cars based on using a label observed through each individual process. The vehicle, ground vehicles, and cars categories are respectively considered superordinate, basic, and subordinate levels. The result of experiment, indicates superordinate advantages. Then, a model based on the human vision system including time, is proposed. The model is then trained for ten categories. The superordinate advantage is also observed in the model results. A possible reason for obtaining these results in the model can be that more neurons are activated in the category at higher levels.

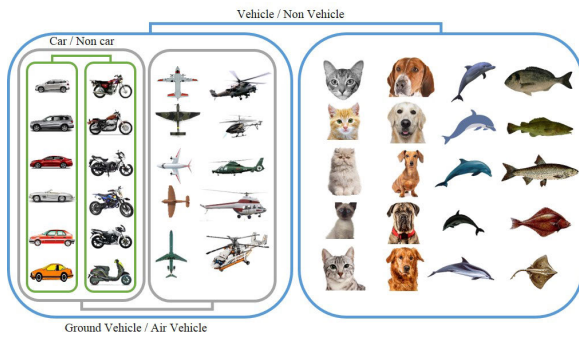
## II. MATERIALS AND METHODS

### A. PARTICIPANTS

In total, ten volunteers, seven males, and three females were included in this experiment. Volunteer age ranged from 20–31 years and mean age 23 years (one left-handed). The ethics committee approved the experiment of the Iran University of Medical Sciences. All participants had normal or corrected to normal vision.

### B. DATASET

We used the images ( $255 \times 255$  pixels, and subtended approximately about  $7^\circ \times 7^\circ$  of visual angle) of eight object categories, including fish, dolphins, cats, dogs, airplanes, helicopters, motorbikes, and cars. Most of the images were selected from the Chandler dataset [23], and others were gathered from some available public sources on the internet. Categorization levels were considered vehicle/ non-vehicle for the superordinate, ground vehicle/ non-ground vehicle for the basic, and car/ non-car for the subordinate level. Figure 1 illustrates several images obtained from the data set for each



**FIGURE 1.** Samples showing some images used in the trial. Three categorization tasks are framed due to stimulus category: blue or the superordinate level, gray for the basic level, and green for the subordinate level. For each categorization level, the target category is illustrated on the left and the distractor category on the right.

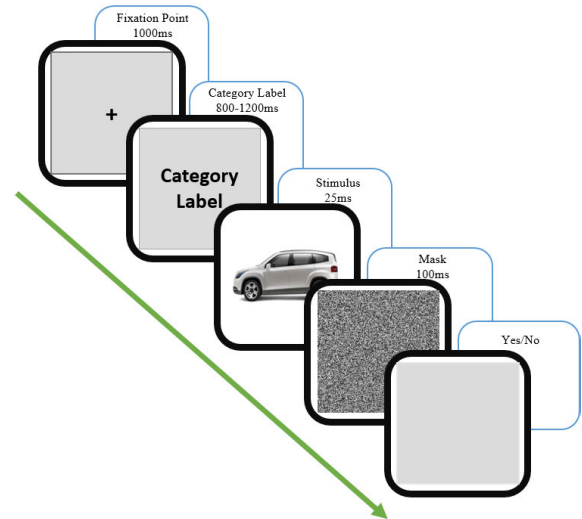
individual task. The distractor consists of everything except the target. The subjects had no prior knowledge concerning the target in the picture.

### C. TASK AND SET-UP

To conduct the experiment, the subjects were seated in a dark room about 50cm away from a computer screen (Intel core 2 duo processor 2.66 GHz, 4 GB RAM, 85 Hz monitor refresh rate), and MATLAB psychophysics toolbox was used [24]–[26] to perform a speeded category verification task. All images were converted to  $256 \times 256$  pixel grayscale and divided into two blocks. Each block contained 146 images (73 animals and 73 vehicles). Each volunteer responds to two different image blocks (10 subjects  $\times$  2 blocks = 20 responded per block), and we obtained 10 different reaction times for each image. Each experiment started with a fixation point to participant centralization for 1000ms, followed immediately by vehicle, ground vehicle, or car category label image for 800 to 1200ms. Then, the test images were displayed for 25ms, next a noise image remained on the screen for 100ms followed by a blank until the participant responded (see Figure 2). Participants respond by pushing a ‘Yes’ key if the label matches the object shown in the stimulus image and a ‘No’ key if it does not. One-third of category verifications were made at the superordinate level (car, motor-bike, airplane, helicopter), one-third were made at the basic level (ground vehicles), and another one-third were made at the subordinate level (car). Note that the category label and the object in the stimulus must be matched in a correct response.

### D. EXPERIMENTAL RESULTS

We observed high accuracy in the result for all three categorization tasks: Median 93% in superordinate, 90% in basic and 89% in the subordinate (see Figure 3B). As it is mentioned previously, we obtain 10 different reaction times (RT) for each individual image. In the experiment, the correct responses with very high (>1000ms) or very low (<200ms) RT are not acceptable. Therefore RTs greater than 1000ms



**FIGURE 2.** Psychophysics categorization task. A fixation point illustrated for 1000ms to focus on the subject. Then, categorization label onset for 800-1200ms. Next, a gray  $256 \times 256$  image named stimulus is flashed for 25ms followed by a noise mask for 100ms. At the end of the experiment, the subjects are asked to respond to whether the presented image matched with the shown label or not using ‘YES’ or ‘NO’ keys on a keyboard.

were eliminated from analysis first. Furthermore, images with more than six correct responses are analyzed, RT for each individual image is computed as the median of RT among all correct responses.

Figure 3A shows the median, within-class, and between-class variance of RT. As it can be seen, levels are separated enough regarding the average and variance of each level. The procedures at the superordinate level are faster than the basic level, and the procedures at the basic level are faster than the subordinate level. In this experiment, we found a superordinate level advantage in the level of categorization.

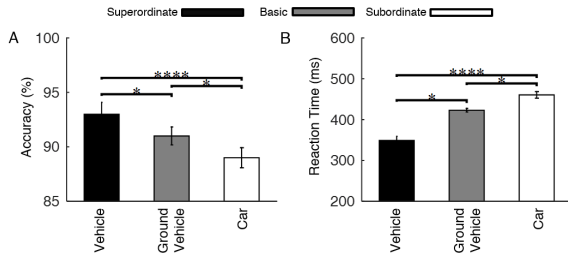
## III. MODEL

### A. BASIC MODEL

In 2007, a novel temporal model based on spike-timing-dependent plasticity (STDP) was introduced [27]. The STDP method is a learning method in the temporal model and involves two important concepts of LTP and LTD. LTP means the previous synapse is activated a few milliseconds earlier than the next synapse, and LTD has the opposite concept. A simplified version of STDP is used in the spiking HMAX model [28].

$$\Delta\omega_{ij} = \begin{cases} a^+ \omega_{ij}(1 - \omega_{ij}), & \text{if } t_j - t_i \leq 0 \\ a^- \omega_{ij}(1 - \omega_{ij}), & \text{if } t_j - t_i > 0 \end{cases} \quad (1)$$

Where  $i$  and  $j$  respectively refer to the index of post and presynaptic neurons,  $t_i$  and  $t_j$  are the corresponding spike times,  $\Delta\omega_{ij}$  is the synaptic weight correction. Besides,  $a^+$  and  $a^-$  are two parameters representing the learning rate. This model, which is based on a spiking network, is looking for stable and discriminative features in the data input. This spike



**FIGURE 3. Accurate than other levels. A) Accuracy for three levels of categorization. It can be seen that superordinate is more accurate than other levels. B) Unlike accuracy, reaction times for the superordinate level are lower than those of another, which means superordinate is the fastest level. These two images show a superordinate advantage.**

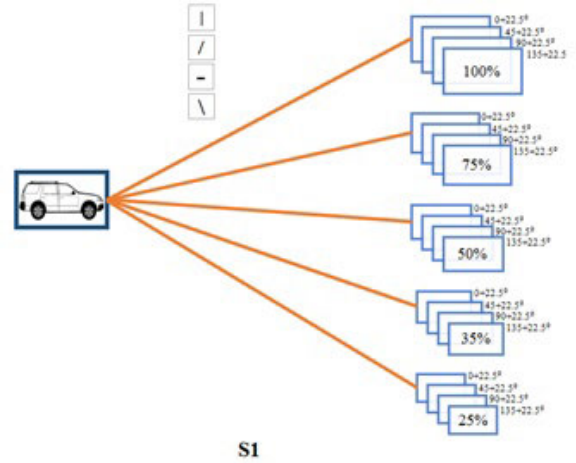
model consists of five layers (S1, C1, S2, C2, and classifier) possessing a similar layout with the HMAX model but a different layout structure. S1 recognizes the edges by performing the convolution a Gabor-Filter on incoming images. This layer is represented as simple cells at the first stage in the visual cortex, as described by Hubel and Wiesel [29]. The cores of Gabor-Filter that is used in this model are  $5 \times 5$  matrices at four angles of  $\pi/8, \pi/4 + \pi/8, \pi/2 + \pi/8, 3\pi/4 + \pi/8$  that the amount of  $\pi/8$  is to avoid focusing on the vertical and horizontal edges. These filters are applied to five different original image sizes (100%, 71%, 50%, 35%, and 25%). Therefore, as it is shown in Fig. 4, there are 20 images as the output of this layer.

C1 layer is similar to a complex vision cortex cell, and each individual C1 complex cell is a local maximum of S1 cells. To calculate the local maximum, a  $7 \times 7$  window moves on the output images of the S1 layer with 6 steps, and then the maximum of that windows is stored as output. Afterward, in the next step, four-images corresponding to four directions in a single size, between their pixel maximization, are carried out. These maximums are saved in a matrix containing coordinate, scale, and orientation, and also a parameter that name is Time Latency. Time Latency is calculated with the following equation:

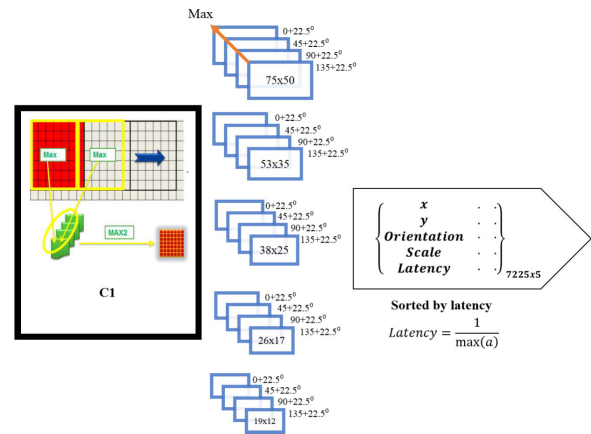
$$\text{Time Latency} = \frac{1}{\max(a)} \quad (2)$$

Where  $\max(\cdot)$  is a maximization operator, and  $a$  is pixels amount of four images. Figure 5 illustrates the C1 layer proceeds.

Integrating these five outputs matrices results in a  $5 \times 7225$  matrix. Time Latency then sorts this matrix. In the next step, data are feeds to S2. S2 unit refers to the mid-level vision features. Each individual neuron is a weight matrix, normally  $16 \times 16$ , in four directions that this weights matrix is initiated randomly filled up with numbers between 0 and 1 and consist 5 matrices with sizes  $75 \times 50, 53 \times 35, 38 \times 25, 26 \times 17$  and  $19 \times 12$ . By feeding the information from the previous layer, the weight is matched with the corresponding neuron, depending on the direction, position and scale. Whenever a cell of neurons reaches a defined threshold (64 based on maximum accuracy), that neuron fires. At this time, the weight



**FIGURE 4. The S1 layer in the entrance of the model. In this layer input image with five scales convolved with four Gabor-Filter in four directions and the outputs of this layer is 20 images [27].**



**FIGURE 5. C1 layer in the model. Maximize operation perform in this layer. The output corresponding to this layer is 4 matrices with the size of  $75 \times 50, 53 \times 35, 38 \times 25, 26 \times 17$  and  $19 \times 12$  which maximize operator performing with a  $7 \times 7$  frame. Then, a maximum operator is performed on all orientation at each scale, and this maximum scale, orientation, and coordinate is stored in a matrix with that latency. The latency is equal to inverted of that maximum value [27].**

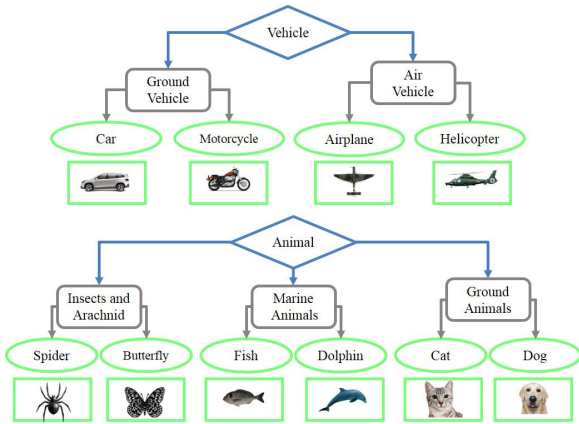
that caused the fire will be updated due to the rule of STDP. At the test phase, all these steps are repeated, but the difference is that when a fire has happened, the numeral of the input spike causing a fire is saved in a spike series matrix. The number of this firing time is called the Firing Rate. Whatever when firing rate grows, the detection time is reduced, and the model recognizes the object faster.

### B. MODEL RESULT

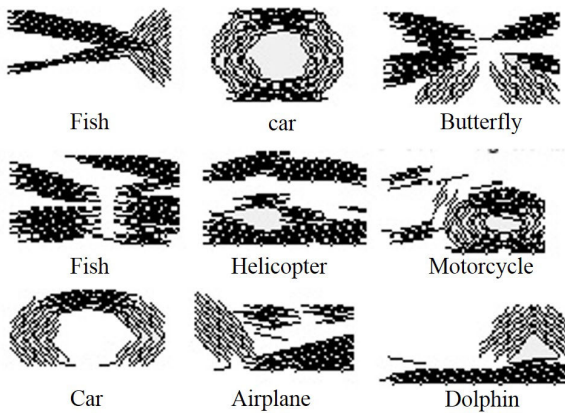
To train the model, ten categories are considered. The selections of these categories are so that they can be seen in different classification levels. Figure 6 shows these categories and levels of classification.

The model is trained using 110 neurons and 60,000 iterations for these 10 categories that each of them contains 100 images. The parameter such as the number of iteration and the number of neurons are selected based on achieving





**FIGURE 6.** Some of the image stimuli and hierarchy of three categorization levels. Different colors determine levels. Blue refers to superordinate, gray highlights the basic, and subordinate is in the green border.

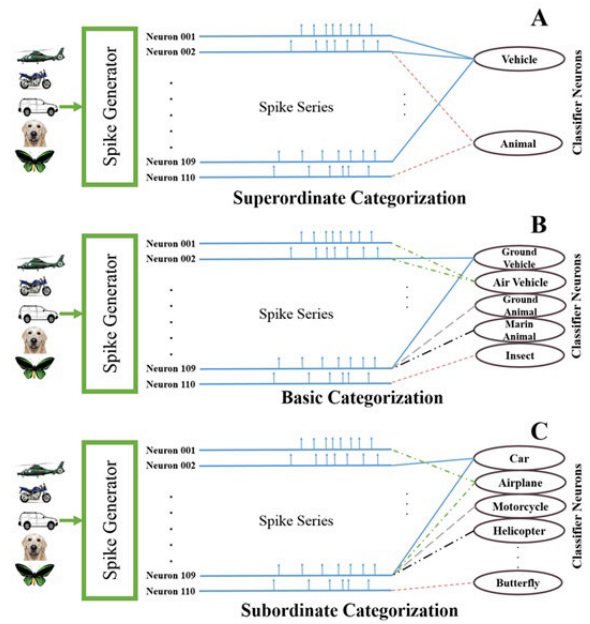


**FIGURE 7.** Final reconstruction of some neurons selectivity to airplane, car, butterfly, fish, helicopter and motorcycle.

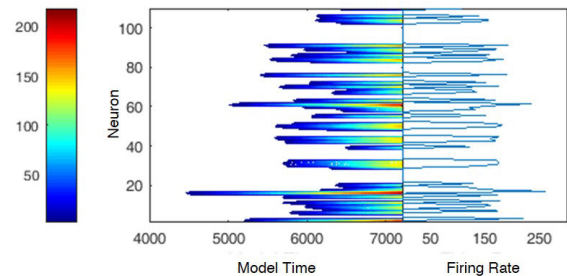
the best accuracy. Upon completion of unsupervised training, different neurons are selected for different input features. Figure 7 shows the trained weights of some neurons for some of the characteristics of the car, airplane, and butterfly.

The classification structure in the model is that the neurons are associated with each class’s properties connected to the single neuron of the next layer of that class. The main reason for choosing this particular type of connection is the effectiveness of level of categorization on classifier connection. Figure 8 illustrates this concept for the three categorization levels.

In fact, if the categorization level is subordinate, it is classifying as a car; only the car-related neurons participate in the categorization. When the aim is to classify ground vehicles, the neurons that describe the characteristics of the car and the motorcycle were selected for classification. In superordinate mode, where categorization aims are vehicles, Neurons related to the car, motorcycle, helicopter, and airplane are in operation. Figure 9 illustrates the firing rate of 110 neurons for cars. With the growth of time, neurons that depend on the car’s features start to fire.

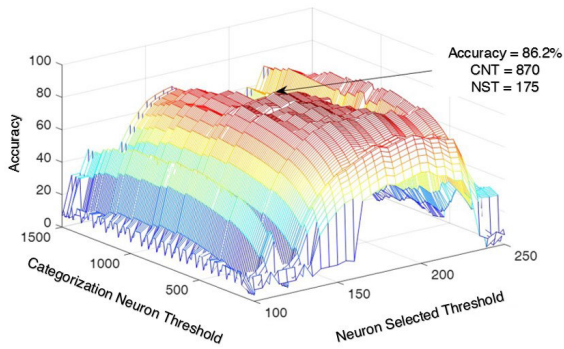


**FIGURE 8.** Three structures of classifier for categorizing three levels of categorization. (A) Categorization structure for the superordinate level. (B) Categorization structure for the basic level (C) Categorization structure for the subordinate level.

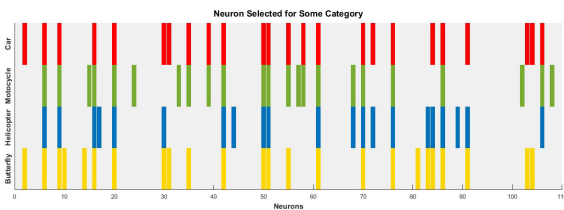


**FIGURE 9.** Car firing rate for one of 110 neurons. Scatter plot is shown the firing rate during the time per neuron. Diagram illustrated the maximum firing rate for each neuron for car input.

The first threshold is used for selection where some neurons are dependent on certain categories, meaning that in the training phase with respect to the initial weights, each of the 110 neurons in the model is dependent on one of ten categories, and to determine which neurons belong to which category, the firing rate of each neuron is examined for the input image of each cluster, and if that firing rate is exceeded to a certain threshold, those neurons will be assigned to that category. The second threshold is used for the neuronal output potential. In fact, as each spike enters from the spike series matrix, the category neuron’s potential for that matrix increases, and when this potential reaches a certain threshold, the input image falls into that category. 10 shows the method for determining these two thresholds. This figure shows the median accuracy of 10 categories in subordinate level for the model in NST and CNT. As can be seen in 10, maximum accuracy is achieved at threshold levels of 175 for



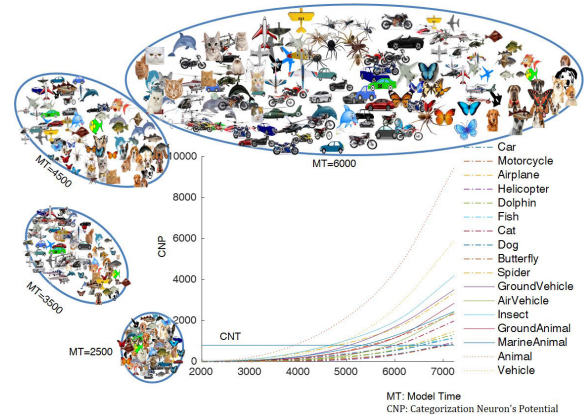
**FIGURE 10.** Median accuracy for 10 categories at subordinate level in CNT and NST. The accuracy of subordinate-level has been used to obtain the best threshold. The best values of these thresholds are calculated in this way which the CNT was changed from 100 to 1500 with steps of 10, and NST was alternately changed from 100 to 250 with 5 steps, and at all these steps, the accuracy of the subordinate level will be counted. The maximum amount of accuracy obtained at CNT = 870 and NST = 175.



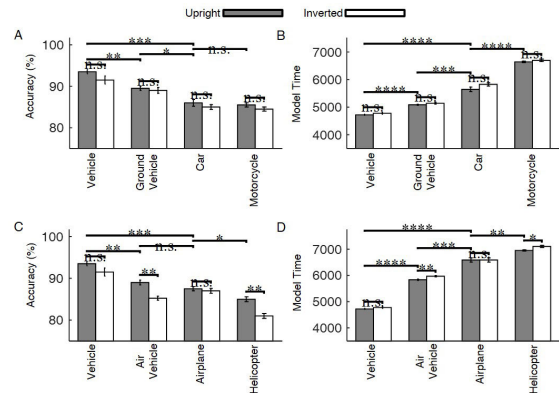
**FIGURE 11.** Neurons that selected for some categories at the 176 threshold values. Car in red, motorcycle in green, helicopter in blue, and butterfly in yellow. These neurons at the training phase are selected to features of each category in form unsupervised mode.

NST and 870 for CNT. The accuracy is 86% in this value of thresholds for subordinate level. As well as, 11 shows some of the neurons of the 110 neurons that belong to certain categories.

At the test stage, the output hardware is specified based on the level of categorization. The categorization is done as follows: For each spike entrance, some amounts are added to the potential of the last layer neuron. In case this potential is reached to a threshold value, the input image is classified into this neuron category. The potential growth in the output neurons is indicated in Figure 12. The categories are started to be separated according to the firing rate of neurons related to each category over time. Finally, we obtain the accuracy and pace of the model for the three levels of categorization. Maximum accuracy is for the superordinate level, and its value is 93%. For all categories at the basic level, the accuracy of ranking is above 89.2%. At the subordinate level, the highest amount of accuracy is related to the spider with 90% and the lowest is for cat, which is 83%. If we compute the median of subordinate level accuracy, it will be 86.2%, which is the lowest efficiency in the categorization levels. In terms of the pace of categorization, superordinate is faster than other levels. Note that in the test phase, images were applied to the model upright and inverse. Figures 13 and 14 show model accuracy and pace for vehicle and animal at three levels of categorization. It can be observed in these charts, accuracy is



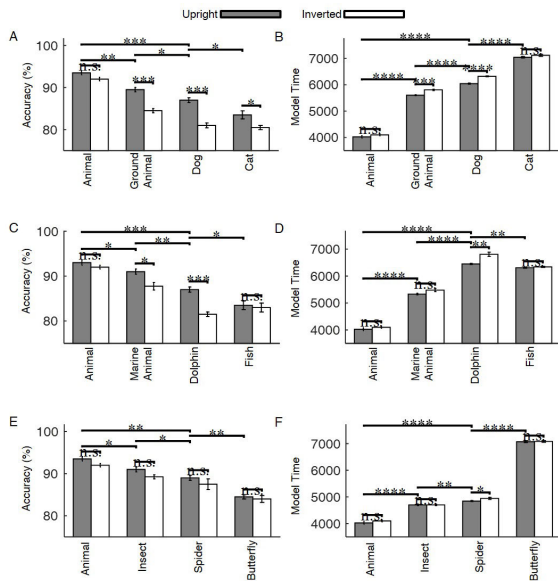
**FIGURE 12.** Potential of output neurons for different data at three levels of classification over time. As time increases, these categories are separated from each other, and the classifier is able to recognize them. The shapes around the chart illustrate detaching this data at four different times. This separation is examined using a K-means classifier with 10 classes. It may be seen that the categorization becomes more precise over time.



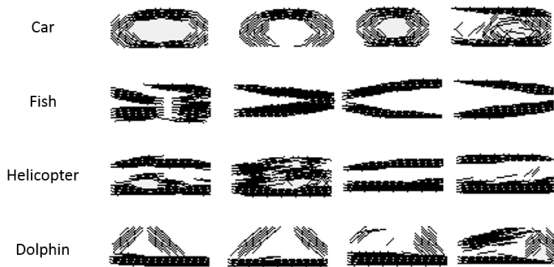
**FIGURE 13.** Accuracy and model time for vehicle input in 10 runs of the model. A) The accuracy of car and motorcycle at upright and inverted images in three levels of categorization. B) Model time of car and motorcycle at upright and inverted images in three levels of categorization. C) The accuracy of airplane and helicopter at upright and inverted images in three levels of categorization. D) Model time of airplane and helicopter at upright and inverted images in three levels of categorization.

very high in all superordinate levels ( $93 \pm 1\%$  for vehicle and  $93.5 \pm 2\%$  for the animal). Due to accuracy and model time, arrangement of categorization is as follows: superordinate, basic and subordinate. These graphs are obtained after 10 times implementing the model with random weights.

Figure 13A shows accuracy and model time for the vehicle, ground vehicle, car, and motorcycle at upright and inverted state. The difference among the three levels is quite significant ( $p < 0.05$ ). However, this has no effect on the arrangement of levels advantages because both of them are at the subordinate level. In figure 13B, the accuracy of basic level and subordinate level at the category of aircraft and helicopter are not significant and the reason is that the number of neurons selected for the airplane is much higher than the helicopter. This is due to the higher diversity of airplanes



**FIGURE 14.** Accuracy and model time for animals input in 10 runs of the model. **A)** The accuracy of dog and cat at upright and inverted images in three levels of categorization. **B)** Model time of dog and cat at upright and inverted images in three levels of categorization. **C)** The accuracy of dolphin and fish at upright and inverted images in three levels of categorization. **D)** Model time of dolphin and fish at upright and inverted images in three levels of categorization. **E)** The accuracy of spider and butterfly at upright and inverted images in three levels of categorization. **F)** Model time of spider and butterfly at upright and inverted images in three levels of categorization.



**FIGURE 15.** Some features related to the categories of car, fish, helicopter and dolphin. These features represent the wheels, and for fish in the car, these features represent dorsal, ventral, and caudal fins. Inverting the images has no effect on these features. However, for the helicopter and dolphin, features are selected to the overall shape of the cabin and the dorsal fin, respectively. Inverted status cannot be represented in these categories.

in comparison with helicopters. However, basic advantage versus subordinate, in terms of the helicopter, can be shown significantly. Figure 14 illustrates the accuracy and the model time for animals in three levels of categorization. For all categories, it can be seen the arrangement of superordinate, basic, and subordinate, respectively. The reason for low and high significance in some categories is that the number of neurons selected for each category is different in the subordinate level.

In the inverse mode, in some categories, there is a sharp drop in precision, and in some cases, there is no change. The reason is that the model is trained for upright images and

certainly for reverse images; the accuracy is reduced, and the diagnostic time also increases. Figure 15 illustrates some features are selected for car, fish, helicopter, and dolphin. As you can see, rotation in some features does not impose any changes.

#### IV. DISCUSSION

In this paper, the issue of levels of categorization has been considered. First, by a rapid psychophysical test, the levels were examined in the human vision system; this experiment's result was a superordinate advantage. Then, a classification model based on the vision system with time was introduced. This model works in that way, for each category, the neurons are selected to that category begin to fire, and in the categorization layer, depending on the level of the category, the neurons associated with that layer are connected to the cluster neuron of the last stage. This model is trained with 110 neurons and 60000 iterations for 100 images per category. Each individual neuron is dependent on some features of data. Then, 50 images were employed for selecting the neuron attributed to each category, and at last, 50 images are used for the test phase. The designed model shows that the superordinate level is faster and more accurate than the other levels. The possible reason for this in the model is that many neurons are involved in the classification at higher levels. Hence, the firing rate in the categorization of higher levels increases, and as a result, the categorization pace goes up.

As noted in the introduction, due to previous studies, there are different opinions about categorization levels. But in some of these studies, it has also been reported the basic level advantages. In fact, the human brain system is a unique hierarchical system and has the same function for categorizing objects. Different factors are the reason for the difference in opinions. Stimulus presentation time, stimulus complexity, and diversity of stimulus are the main reasons for these differences. Poncet and Thorpe in 2014 are decisive have stated superordinate level advantage. They carried out several experiments, in which experiments change the flashing time of stimulus and change the type of categorization were visible [17]. Mack and Palmeri in 2015 were rejected the theory of Poncet and Thorpe. They introduced superordinate advantages in rapid categorization same as Mack and Palmeri but they explained another result in long flashed stimuli. They were designed two experiments. In the first experiment, the stimulus presentation was only 25ms, and the mask was presented 250ms after that and in the second experiment, the stimulus was presented 250ms, followed by a mask. It was actually considered, respectively superordinate advantage at the first experiment and a basic advantage in the second experiment. They attributed this to the fact that the coding time of the information was separate and independent from the time of categorization. Moreover, superordinate advantages were presented in the explained model in this research that associated a rapid categorization experiment, and it is because of changes in the function of the brain in the change of category type. When category type is changed,



the hardware of the brain also changed, and IT area is selected the different number of neurons for categorization.

Another important look at this issue is the size of the category area. At the higher level, the inter-class area becomes larger, and therefore the classification becomes simpler. In fact, according to this that how enormous is an intra-class domain, the speed of categorization can be affected [19].

At higher levels, more areas are included in the categorization, which was introduced by Bamer and Just in 2017 [30]. They examined two levels of basic and subordinate and expressed the basic level advantages. In their opinion, basics has the advantage because humans classify more basic levels in their life and are experts in basic classification. As it is shown in the test phase, the accuracy is reduced and detection time is increased by applying images in inverted mode, and the decrease is almost the same for all three levels. In fact, the reason for this decline is that the model is trained with upright images. At the time of training, the model learns some features of each category. If those features are changed with inverted images in the test phase with inverted images, accuracy decreases sharply. However, if those features in inverted images to be presented, accuracy is changed gradually. For example, most features are wheels, and the wheel does not change with the car category's inversion. Hence, car accuracy and model time decreases very low and those neurons are to wheel also fire. These changes can be seen in the same way at higher levels. As is seen in Figure 15, there are no significant changes in reverse mode for the car and fish features. Whereas, the associated feature is lost in reverse mode for helicopters and dolphins. Therefore, the type of categories determines their accuracy for the inverted state, and this experiment cannot be determined the category pathway for each level.

## REFERENCES

- [1] J. Martinovic, T. Gruber, and M. M. Müller, "Coding of visual object features and feature conjunctions in the human brain," *PLoS ONE*, vol. 3, no. 11, Nov. 2008, Art. no. e3781.
- [2] S. Thorpe, D. Fize, and C. Marlot, "Speed of processing in the human visual system," *Nature*, vol. 381, no. 6582, pp. 520–522, Jun. 1996.
- [3] M. Fabre-Thorpe, G. Richard, and S. J. Thorpe, "Rapid categorization of natural images by rhesus monkeys," *NeuroReport*, vol. 9, no. 2, pp. 303–308, Jan. 1998.
- [4] A. Mirzaei, S.-M. Khaligh-Razavi, M. Ghodrati, S. Zabbah, and R. Ebrahimpour, "Predicting the human reaction time based on natural image statistics in a rapid categorization task," *Vis. Res.*, vol. 81, pp. 36–44, Apr. 2013.
- [5] H. Eichenbaum, "Barlow versus Hebb: When is it time to abandon the notion of feature detectors and adopt the cell assembly as the unit of cognition?" *Neurosci. Lett.*, vol. 680, pp. 88–93, Jul. 2018.
- [6] A. Delorme, G. Richard, and M. Fabre-Thorpe, "Ultra-rapid categorisation of natural scenes does not rely on colour cues: A study in monkeys and humans," *Vis. Res.*, vol. 40, no. 16, pp. 2187–2200, Jul. 2000.
- [7] J. Antolík, C. Monier, Y. Frégnac, and A. P. Davison, "A comprehensive data-driven model of cat primary visual cortex," *BioRxiv*, Jan. 2019, Art. no. 416156. [Online]. Available: <https://www.biorxiv.org/content/early/2019/11/08/416156>, doi: 10.1101/416156.
- [8] D. J. Freedman, "Categorical representation of visual stimuli in the primate prefrontal cortex," *Science*, vol. 291, no. 5502, pp. 312–316, Jan. 2001.
- [9] G. A. Lean, Y.-J. Liu, and D. C. Lyon, "Cell type specific tracing of the subcortical input to primary visual cortex from the basal forebrain," *J. Comp. Neurol.*, vol. 527, no. 3, pp. 589–599, Feb. 2019.
- [10] S. Zabbah, K. Rajaei, A. Mirzaei, R. Ebrahimpour, and S.-M. Khaligh-Razavi, "The impact of the lateral geniculate nucleus and corticogeniculate interactions on efficient coding and higher-order visual object processing," *Vis. Res.*, vol. 101, pp. 82–93, Aug. 2014.
- [11] H. Kim, G. Kim, and S.-H. Lee, "Effects of individuation and categorization on face representations in the visual cortex," *Neurosci. Lett.*, vol. 708, Aug. 2019, Art. no. 134344.
- [12] T. Brinda, S. Napierala, D. Tobinski, and I. Diethelm, "Student strategies for categorizing IT-related terms," *Educ. Inf. Technol.*, vol. 24, no. 3, pp. 2095–2125, May 2019.
- [13] T. T. Rogers and K. Patterson, "Object categorization: Reversals and explanations of the basic-level advantage.," *J. Exp. Psychol., Gen.*, vol. 136, no. 3, pp. 451–469, Aug. 2007.
- [14] M. Poncet and M. Fabre-Thorpe, "Stimulus duration and diversity do not reverse the advantage for superordinate-level representations: The animal is seen before the bird," *Eur. J. Neurosci.*, vol. 39, no. 9, pp. 1508–1516, May 2014.
- [15] M. W. Morris and G. L. Murphy, "Converging operations on a basic level in event taxonomies," *Memory Cognition*, vol. 18, no. 4, pp. 407–418, Jul. 1990.
- [16] A.-H. Vahabie, M.-R. A. Dehaqani, M. N. Ahmadabadi, B. N. Araabi, and H. Esteky, "Rapid face adaptation distributes representation in inferior-temporal cortex across time and neuronal dimensions," *Sci. Rep.*, vol. 7, no. 1, p. 1709, 2017.
- [17] E. Rosch, C. B. Mervis, W. D. Gray, D. M. Johnson, and P. Boyes-Braem, "Basic objects in natural categories," *Cognit. Psychol.*, vol. 8, no. 3, pp. 382–439, Jul. 1976.
- [18] M. L. Mack, I. Gauthier, J. Sadr, and T. J. Palmeri, "Object detection and basic-level categorization: Sometimes you know it is there before you know what it is," *Psychonomic Bull. Rev.*, vol. 15, no. 1, pp. 28–35, Feb. 2008.
- [19] K. Mohan and S. P. Arun, "Similarity relations in visual search predict rapid visual categorization," *J. Vis.*, vol. 12, no. 11, p. 19, Oct. 2012.
- [20] M. L. Mack and T. J. Palmeri, "The dynamics of categorization: Unraveling rapid categorization," *J. Exp. Psychol., Gen.*, vol. 144, no. 3, p. 551, 2015.
- [21] M.-R.-A. Dehaqani, A.-H. Vahabie, R. Kiani, M. N. Ahmadabadi, B. N. Araabi, and H. Esteky, "Temporal dynamics of visual category representation in the macaque inferior temporal cortex," *J. Neurophysiol.*, vol. 116, no. 2, pp. 587–601, Aug. 2016.
- [22] M. N. Ashtiani, S. R. Kheradpisheh, T. Masquelier, and M. Ganjtabesh, "Object categorization in finer levels relies more on higher spatial frequencies and takes longer," *Frontiers Psychol.*, vol. 8, p. 1261, 2017.
- [23] D. M. Chandler and D. J. Field, "Estimates of the information content and dimensionality of natural scenes from proximity distributions," *J. Opt. Soc. Amer. A, Opt. Image Sci.*, vol. 24, no. 4, p. 922, Apr. 2007.
- [24] D. H. Brainard, "The Psychophysics Toolbox," *Spatial Vis.*, vol. 10, no. 4, pp. 433–436, 1997.
- [25] M. Kleiner, D. Brainard, D. Pelli, A. Ingling, R. Murray, and C. Broussard, "What's new in psychtoolbox-3," *Perception*, vol. 36, no. 14, pp. 1–16, 2007.
- [26] D. G. Pelli, "The videotoolbox software for visual psychophysics: Transforming numbers into movies," *Spatial Vis.*, vol. 10, no. 4, pp. 437–442, 1997.
- [27] T. Masquelier and S. J. Thorpe, "Unsupervised learning of visual features through spike timing dependent plasticity," *PLoS Comput. Biol.*, vol. 6, Feb. 2007, Art. no. e31.
- [28] H. H. Gorji, S. Zabbah, and R. Ebrahimpour, "A temporal neural network model for object recognition using a biologically plausible decision making layer," Jun. 2018, *arXiv:1806.09334*. [Online]. Available: <https://arxiv.org/abs/1806.09334>
- [29] D. H. Hubel and T. N. Wiesel, "Receptive fields, binocular interaction and functional architecture in the cat's visual cortex," *J. Physiol.*, vol. 160, no. 1, pp. 106–154, Jan. 1962.
- [30] A. J. Bauer and M. A. Just, "A brain-based account of 'basic-level' concepts," *NeuroImage*, vol. 161, pp. 196–205, Nov. 2017.





**MOHAMMAD HOSSEIN KARIMI** was born in Yazd, Iran, in 1987. He received the B.S. degree in electronics engineering from Shahid Rajaei Teacher Training University, in 2010, and the M.S. degrees in electronics engineering from the Khaje Nasir Toosi University of Technology, in 2013. He is currently pursuing the Ph.D. degree in electronics engineering. His research interests include the neuroscience, object categorization, and brain modeling.



**REZA EBRAHIMPOUR** received the Ph.D. degree in cognitive neuroscience from the SCS, IPM, Tehran, Iran, in July 2007. He is currently a Professor with the Department of Computer Engineering, Shahid Rajaei Teacher Training University, Tehran. He is the author or coauthor of more than 100 international journals and conference publications in his research areas, which include *Cognitive and Systems Neuroscience*, *Human and Machine Vision*, *Decision Making*, and *Object Recognition*.



**NASOUR BAGHERI** received the M.S. and Ph.D. degrees in electrical engineering from the Iran University of Science and Technology (IUST), Tehran, Iran, in 2002 and 2010, respectively. He is currently an Associate Professor with the Electrical Engineering Department, Shahid Rajaei Teacher Training University, Tehran. He is the author of over 100 articles in information security and computational neuroscience. His research interests include cryptology, more precisely, designing and analysis of symmetric schemes, such as lightweight ciphers, e.g., block ciphers, hash functions, and authenticated encryption schemes, cryptographic protocols for constrained environment, such as RFID tags and the IoT edge devices and hardware security, e.g., the security of symmetric schemes against side-channel attacks, such as fault injection and power analysis. In the field of neuroscience, he is more interested in the computational modeling of human vision system. His web-page is: <https://sites.google.com/view/nasour-bagheri>.

...