Do Deep Learning Models Mimic Human Personality Traits? – An Empirical Study

Anirudh A
Department of Computer Science and Information Systems
Birla Institute of Technology and Science (BITS), Pilani
Hyderabad, India
Email: f20180936@hyderabad.bits-pilani.ac.in

Subrata Chakraborty
Faculty of Science, Agriculture, Business and law, University of New England, Armidale, NSW, 2531, Australia
Centre for Advanced Modelling and Geospatial Information Systems (CAMGIS), Faculty of Engineering and IT, University of Technology Sydney, Sydney, NSW, 2007, Australia
Email: Subrata.chakraborty@une.edu.au

Abstract

Two key aspects of artificial intelligence are its ability to make decisions and attempt to mimic humans. Decision-making in humans is, however not straightforward and depends significantly on the person's mental state, personal biases, and personality. In this study, we attempt to empirically understand if deep learning image classifiers also exhibit such inherent biases or if they act neutrally in any given situation. To this end, we perform three experiments – left-brain right-brain test, psychological images test, and Rorschach's inkblot test on eight different stat-of-the-art deep learning classifiers. A detailed analysis of the SoftMax probability scores is done rather than an analysis on measures like accuracy and F1. The experimental results suggested that most models work similar to a left-brained person, do not always predict the same class when given images consisting of multiple object classes, and usually detect larger objects rather than smaller ones. We believe that understanding these inherent biases would help future researchers take necessary actions while building image classification models.

Keywords: Psychological image test, lateralization of brain function, data augmentation, Rorschach's inkblot test, image classification.

1. Introduction

In several day-to-day situations, amidst all the clutter, humans pick up and identify things they like, fear, or are attached to in some fashion first. Similarly, while reading a book or watching a movie, one connects with aspects closer to themselves than others. In technical terms, human decisions are often biased based on personality traits and brain functionality. One such decision-making process is recognizing objects amongst several other objects, and several tests exist that study one's personality traits based on what one sees first in an image (Sutherland et al. 2015). In this study, we test whether deep learning models show such bias traits through an empirical study by conducting three image classification experiments.

At the most basic level, the human decision-making process can be considered a task of choosing between two courses of actions in order to attain a final goal. It is an outcome of careful evaluation of alternative options concerning their likelihood and value of outcomes. However, it has been observed that the human decision-making process is not as straightforward as it seems – emotions are critical in decision making, past decisions tend to affect future ones, and the process is strongly biased by unconscious mental processes (Power et al. 2011). While decisions can be categorized as strategic, tactical, and operational, from the psychological perspective, people tend to adopt one of the two decision-making strategies – (i) availability heuristics, and (ii) representative heuristics. While the availability aspects are based on how easily one can remember similar events from the past, the representative aspect involves comparing the current situation with a prototype of a particular event.

The human brain is divided into two hemispheres – left and right. The concept of brain function lateralization (Hugdahl 2000) states that cognitive processes are specialized to one side of the brain. While the left brain is used for critical thinking, reasoning, logical thinking, language, and number understanding, etc., the right brain deals with aspects such as arts, intuition, imagination, emotions, sexual preferences, facial recognition, etc. A systematic deviation from this rationality in judgment is known as the cognitive bias (Haselton et al. 2015). One of the aspects of decision-making is the process of image recognition and classification. Although vision seems straightforward and trivial from a superficial understanding, research work suggests the brain processes more than 60 images, each with millions of pixels every second. Similarly, the psychological image tests (Sutherland et al. 2015) attempt to understand various aspects of a person's personality based on what they recognize first from a given image consisting of multiple classes of objects, and these tests also enable psychologists to understand the person's mental state.

Information systems that support and organize decision-making activities are known as decision support systems (DSS) (Bonini 1963). They are computerized programs that support judgement, determination, and courses of action. With artificial intelligence and deep learning techniques performing exceedingly well in several tasks, intelligent decision support systems (IDDSs) (Ahmad and Simonovic 2006) have started gaining popularity. The applications of IDDSs can be seen in various fields such as diabetes prediction (Yahyaoui et al. 2019), brain tumor classification (Sharif et al. 2021), product recommendations in eCommerce sites (Acharjee et al. 2017), agriculture (Kukar et al. 2017), real estate (Del et al. 2019), etc.

Deep learning techniques have been found to perform exceedingly well in computer vision – image classification and object detection in particular and have also gained importance in decision support systems. Considering the fact that a key aspect of artificial intelligence is how closely machines can mimic humans, but at the same time, biases in the decision-making process are not desired, we perform an empirical study to understand how eight state-of-the-art deep learning classifiers – MobileVNet (Howard et al. 2017), Single Shot Detector (SSD) (Liu et al. 2016), AlexNet (Krizhevsky et al. 2012), InceptionV3 (Szegedy et al. 2016), VGG19 (Simonyan and Zisserman 2014), ResNet50 (He et al. 2016), GoogleNet (Szegedy et al. 2015), and DenseNet121 (Huang et al. 2017) react to three psychological image tests – the left-brain-right-brain test (Denny and Wolf 1984), the personality (or psychological) image test (Sutherland et al. 2015), and the Rorschach Inkblot test (Vernon 1933; Hubbard and Hegarty 2016). These experiments are conducted to understand if there exists any inherent bias in the models and further see which human personality traits (if any) they mimic. We believe that understanding the inherent biases would help future researchers take necessary actions while building image classification models.

2. Related Work

The concept of lateralization (Güntürkün et al. 2019) of the brain and its functions suggests that certain neural functions tend to be more specialized towards one side of the brain than the other. Brain asymmetries (Iaccino 2014) are key components of humans' and other animals' cognitive, sensory, and motor systems. Studies in this area date back to as early as 1865 (Broca 1865), when it was first seen that the brain's left hemisphere is involved in speech among humans.

The psychological and emotional state of a person tends to have an effect on their perception of colour (Mikellides 2012). While the left brain perceives light colours easily, the right brain is sensitive towards the darker ones. In fact, the laterality patterns of brain functional connectivity are also said to have a gender-based effect (Ingalhalikar et al. 2014). It is also suggested that the most obvious sign that the brain functions asymmetrically is that majority of humans tend to prefer the right hand over the left (Corballis 2014). Testing of each of the disconnected hemispheres has revealed that the left brain specializes in language and logical thinking while the right brain is used for emotional and nonverbal functions. The personality, cognitive style, and decision-making ability of a person depend on which hemisphere of the brain dominates, i.e., is the person left-brained or right-brained. A left-brained person is usually more logical, well-organized, and inclined towards non-fiction, while a right-brained person is said to be more creative, emotional, artistic, and spontaneous (MacNeilage et al. 2009).

It has been claimed that the first thing one sees in a picture could reveal their personality, state of mind, and character. Such psychological image tests could use drawings, double-sided snapshots, and abstract photographs to understand the test-taker's mental state and personality. It is however essential to pay attention to the first object that one sees since one would definitely identify multiple images after a while. One such projective and subjective test that is widely used by psychologists is the Thematic Appreciation Test (TAT) (Nissley and DeFreese 2020) which involves the subject describing ambiguous scenes. Yet another test that uses a very similar approach is the Rorschach inkblot test (Vernon 1933; Hubbard and Hegarty 2016). With a reliability of around 89% (Hertz 1935), the inkblot test records the subjects' perceptions of inkblots (abstract images) and analyzes the results using complex algorithms and psychological interpretation. Like other psychological tests, the inkblot test also tries to examine the subject's emotional functioning and personality with a focus of determining thought disorder, particularly in situations where the patient is unwilling to describe their thought process openly.

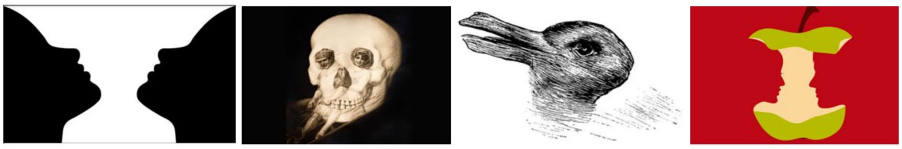
On the other hand, deep learning classifiers have started gaining great popularity in tasks such as image classification and object detection. An exhaustive review of various state-of-the-art neural network-based image classifiers is done in Chen et al. 2021. The paper discusses the basic structure of artificial neural networks and convolutional neural networks, studies the various classifiers in detail and performs a comparative analysis on the chosen techniques. He et al. 2019 suggest a collection of refinements apart from those related to the training procedures to improve the performance of the existing models such as ResNet50.

A significant amount of work has been done in the broad field of optical illusions detection (Gomez-Villa et al. 2019; Gomez-Villa et al. 2020), which conclusively prove that AI and CNNs are deceived by optical illusions just like humans. Models trained on motion videos get deceived and recognise motion-like images (illusions) to be videos as well and, in fact, predict the direction of motion too. Similarly, CNNs trained on natural images tend to reach just like humans when tested on visual (colour-based) optical illusions. Whether CNNs should actually reproduce optical illusions or not has, however, been a topic of debate (Lonnqvist et al. 2021). The key takeaway, however, has been that one could make much better use of CNNs by focussing more on their differences rather than similarities with humans.

Despite work being done in the field of multi-object detection (Pal et al. 2021; Wang et al. 2018; Wang et al. 2021), and psychological testing, to the best of our knowledge, no work has explored how image classifiers react when multiple classes of objects are present in the same image and how deep learning neural networks react to tests like the left-brain right-brain test, psychological image test, and the inkblot test. This study explores the results of such tests (which are usually performed on humans) on several deep learning models through an empirical testing.

3. Dataset Description

In this work, three experiments (left-brain right-brain test, personality test, inkblot test) are performed on each classifier. Each test image consists of two to five classes of objects. All models have been trained on the classes present in the test images by scraping 300 images from Google and removing any irrelevant image from the dataset manually. The left-brain right-brain test is performed on eight test images, samples of which are shown in Figure 1. For the test images A-D in Figure 1, the two classes of objects are faces and flower vase, skull and children, duck and rabbit, and apple and faces, respectively. The personality tests have been performed on thirty images, samples of which are shown in Figure 2, with each image trying to address one specific personality aspect. For the test images A-D in Figure 2, the two classes of objects are a skull and two astronauts, binoculars and a car, an infant, a couple, and a tree, and trees, birds, and a hut, respectively. Finally, the inkblot test is performed on ten test images (samples of which are shown in Figure 3) that have been originally suggested in Rorschach's psychological publication (Hubbard and Hegarty 2016). The Rorschach's test has been conducted on the benchmarked dataset published while curating the experiment (Vernon 1933). However, since left-brain right-brain test and personality tests are in their early stages of study, to the best of our knowledge, no such benchmarked dataset is available, and hence the test images have been manually scraped from Google (Times Now Personality Tests).



(A) (B) (C) (D)

Figure 1: Test Images for left-brain right-brain test

(A) (B) (C) (D)

Figure 2: Test Images (Samples) for Psychological Image Test

(A) (B) (C) (D)

Figure 3: Test Images for Rorschach Inkblot Test

4. Experimental Setup and Study Design

In this work, three experiments have been performed on various image classification techniques to understand which humanistic traits (if any) they try to mimic and to see if the classification techniques have a bias towards any specific class of images. Eight state-of-the-art deep learning classifiers – MobileVNet (Howard et al. 2017), Single Shot Detector (SSD) (Liu et al. 2016), AlexNet (Krizhevsky et al. 2012), InceptionV3 (Szegedy et al. 2016), VGG19 (Simonyan and Zisserman 2014), ResNet50 (He et al. 2016), GoogleNet (Szegedy et al. 2015), and DenseNet121 (Huang et al. 2017) are studied in the three experiments. To maintain uniformity, all the models have been trained on the ImageNet data, compiled using sparse categorical cross entropy function and the Adam optimizer. In each model, the SoftMax layer has been used to make the final predictions. A learning rate scheduler (Li et al. 2019) is used to train the models with multiple training rates (0.01, 0.001, 0.0001, and 0.00001 – picked based on trail and error) at different epochs. Image pre-processing includes data augmentation (Wong et al. 2016) (rotation, random zoom, width/height shift, and horizontal/vertical flip) performed before training to enhance the predictive ability of the models. The summary of the three experiments is depicted pictorially in Figure 4, and the following sub-sections explain each experiment in detail:

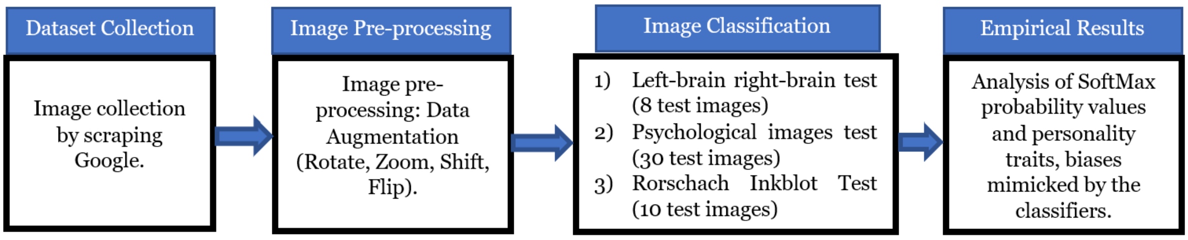


Figure 4: Experimental Setup

4.1 Experiment 1: Left-Brain Right-Brain Test

The human brain is said to be divided into two halves (or hemispheres) – left and right. Studies (Hugdahl 2000) suggest that some neural or cognitive functions tend to be dominated by one of the two hemispheres. In general, every person is said to be either left-brain or right-brain dominated. To understand if a person is left-brained or right-brained, three types of tests are usually conducted – (i) colour related tests – images where there exists ambiguity in colour are shown to the subject (left-brained people tend to recognize lighter colours while right brained people are sensitive to darker colours (Zeki and Marini 1998)), (ii) motion related tests – the right brain is usually responsible for recognising clockwise motion while the left brain is associated with anti-clockwise motion, and (iii) confusing image tests – images where multiple classes are merged are shown to the subject and analysis is performed based on what class they recognise first. Since deep learning techniques use numerical red-green-blue (RGB) values, it is not possible to analyse them based on the perception of colour, and at the same time, image classification techniques cannot be used to study the direction of motion in videos. Hence, in this study, each image classification model is tested on images where each test image is considered a binary class classification problem. Samples of the eight test images used in this experiment are shown in Figure 1. A majority vote is taken based on the results of each test image in order to make the final conclusion.

4.2 Experiment 2: Psychological Images Test

Psychological image tests are conducted to analyse various personality traits of the test-taker. Unlike the left-brain right-brain test, the psychological image tests usually do not involve confusing images. They use images where two or more classes of objects can be clearly distinguished from one another. Depending on what object the test-taker spots first, various traits of his personality are analysed. In this work, each image classification model is tested on thirty such images, examples of which are shown in Figure 2 to understand which human personality (if any) each model tends to mimic.

4.3 Experiment 3: Rorschach's Inkblot Test

An inkblot test (Vernon 1933), published by Swiss psychiatrist Herman Rorschach is yet another personality test that attempts to study the personality of the test taker by evaluating their response to ambiguous inkblots. When the test is used empirically, the measurement of personality (how normal/abnormal a person is) is related to the quality of the responses. Despite the test being around 100 years old, the experiments described in its authentic form are still relevant today (Hubbard and Hegarty 2016). Since creative thought process in itself depends on a person's mental state and personality, we've chosen to perform the Rorschach's Inkblot Test in addition to the first two experiments.

In this work, the ink blot test is performed on the eight deep learning classifiers by considering each test image as a three, four, or five class image classification problem. Based on the predictions on all test images (Figure 3), the personality traits of each image classification technique are understood. Since the test images of all three experiments have multiple classes of objects, all of which are possible predictions, measures like accuracy, F1-score, etc., are not used in this work. Instead, the probability values obtained from the SoftMax function of the output layer from each model are analysed. Each model is run ten times on each test image, the mean and median are computed for the probability scores for each class across these ten iterations, and a majority vote is taken which is considered as the model's final prediction. This is done because the models tend to predict the classes with different probabilities each time since the images are confusing and the classes are not mutually exclusive in each image.

5. Results and Discussion

5.1 Results from Experiment 1

The goal of this experiment was to understand if each of the eight deep learning classifiers are similar to left-brained or right-brained people or if they act neutrally. The models used in this study are trained and evaluated on ImageNet, consisting of 1000 classes. All the test images in this experiment are considered as two-class classification problems. The models were run ten times on each test image, and the mean, median, and standard deviation of the SoftMax probability scores were computed. Of the ten iterations for each image, a majority vote was taken to determine if each classifier acted left-brained or right-brained for that particular test image. Further, a majority vote was again taken considering the results of the eight test images to see if the classifiers acted left-brained or right-brained in general. Table 1 shows the SoftMax probability scores for the ten trials conducted on image 'A' in Figure 1. With respect to image 'A' in Figure 1, it can be seen that AlexNet and MobileVNet act similar to a right-brained person, while other models act similar to a left-brained person. Further probability values very close to 0.5 for models such as VGG19, SSD, DenseNet, etc., suggest that the models were able to recognise both classes of images considerably in each iteration, while models such as MobileVNet and GoogleNet give probability scores very close to 1 (and 0 for the other class) in most cases. Since models with alternating high and low probabilities for a particular class could also give a mean value of around 0.5, the majority vote is taken rather than considering the mean probability score directly. It can also be seen that models such as ResNet50 and AlexNet do not make the same prediction of the recognised class each time. Quantitatively, these are, in fact, the models which have the highest standard deviation with respect to the probability values. This is similar to the observation made by Gomez et al. 2020 which states that deep learning models, like humans, are deceived by optical illusions.

	Class	Mobile V Net	VGG 19	Dense Net	Inception V3	ResNe t 50	SSD	Alex Net	Google Net
Mean	Vase	0.028	0.580	0.662	0.592	0.578	0.556	0.393	0.742
	Face	0.972	0.420	0.338	0.408	0.422	0.444	0.607	0.258
Medi-an	Vase	0.027	0.573	0.626	0.649	0.606	0.512	0.141	0.993
	Face	0.973	0.427	0.374	0.351	0.394	0.488	0.859	0.007
Std Dev	Vase	0.020	0.139	0.140	0.212	0.314	0.157	0.446	0.362
	Face	0.020	0.139	0.140	0.212	0.314	0.157	0.446	0.362
Majority Vote		Faces	Vase	Vase	Vase	Vase	Vase	Faces	Vase
Conclusion		right	left	left	left	left	left	right	left

Table 1: Results from Experiment 1 (Left-Brain Right-Brain Test)

The results corresponding to image 'A' in Figure 1 are summarised using box plots in Figure 5. From the plot, it is evident that all probability scores of SSD are very close to 0.5. The low standard deviation and a mean value close to 0 (or 1) for MobileVNet suggests that the model was highly confident in its predictions and predicted the class to be 'Faces' in all ten iterations. A plot spanning a larger area for AlexNet and GoogleNet indicates that the models have been predicting both classes of images almost equally. Table 2 shows the majority vote results for each of the eight test images. The final conclusion is made by taking a majority vote from these eight results. Most models (MobileVNet, DenseNet, InceptionV3, ResNet50, AlexNet, and GoogleNet) act similar to a left-brained person while SSD acts neutrally and VGG19 acts similar to a right-brained person. The results make sense logically since the left brain in humans is responsible for critical thinking, reasoning, logic, math skills, etc., which deep learning classifiers are built to perform.

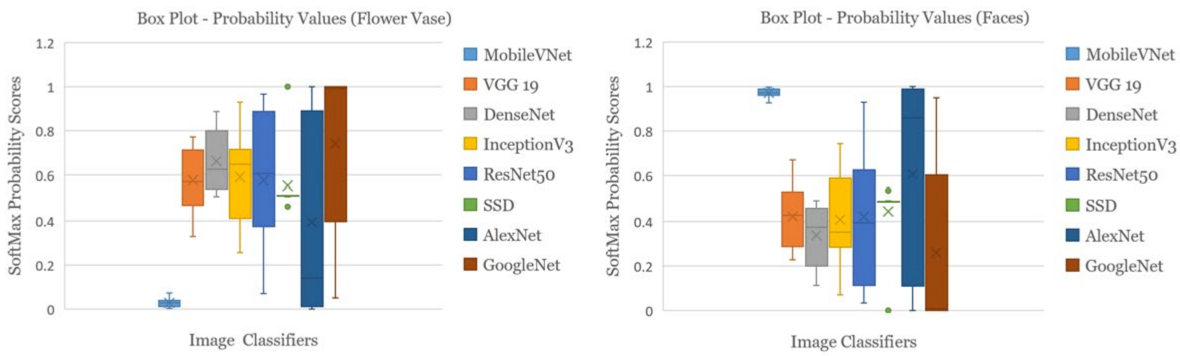


Figure 5: Box Plot for Probability Values Corresponding to Image 'A' in Figure 1

Model	Image 1	Image 2	Image 3	Image 4	Image 5	Image 6	Image 7	Image 8	Conclusion
MobileVNet	right	right	left	left	left	left	left	left	left
VGG 19	left	right	right	right	right	right	left	left	right
DenseNet	left	left	tie	left	right	left	right	tie	left
InceptionV3	left	left	right	right	left	left	left	left	left
ResNet50	left	left	right	left	left	tie	right	left	left
SSD	left	right	right	left	right	right	left	left	neutral
AlexNet	right	left	right	left	left	left	tie	left	left
GoogleNet	left	left	right	left	left	left	right	left	left

Table 2: Majority Vote Results - Experiment 1

5.2 Results from Experiment 2

In this experiment, we attempted to understand if the deep learning classifiers mimicked any particular human personality trait. Each model was tested on 30 pictures (samples of which are shown in Figure 2), each containing two to five classes (each class representing a particular personality trait). The final predictions were made based on a majority vote similar to the previous experiment.

Table 3 describes the results pertaining to image ‘C’ in Figure 2. A high standard deviation value and a mean value of around 0.5 for MobileVNet and VGG19 suggest that they recognise both classes of images equally unlike InceptionV3 and SSD. Probability values far from 0.5 for a binary classification problem indicate that the models do not make a neutral and unbiased prediction despite all classes of objects being present in the image. Taking a closer look at the experimental results also suggests that when one of the classes in the image is very small in size as compared to the other class, the probability of it being recognised in the presence of the other is very small and close to zero. One such example is shown in image ‘G’ of Figure 2 where the sketch of birds is very small as compared to that of the tree and the hut and the corresponding results are shown in Table 4. This suggests that future researchers should take special care while using these models to detect tiny objects apart from taking measures to account for the inbuilt biases. Finally, Table 5 enlists some of the personality traits that each image classification model mimics. These results are summarised based on the OCEAN model of personality traits where O, C, E, A, and N stand for Openness, Conscientiousness, Extraversion, Agreeableness, and Neuroticism, respectively. Although personality traits may not make a lot of sense with respect to deep learning models and machines, a non-neutral personality could mean that the models have some inbuilt biases towards one particular class. Despite there being debates on whether CNNs should mimic human personalities in such a fashion (Lonnqvist et al. 2021), this experiment conclusively proves that CNNs actually mimic personality traits and are deceived by illusion-like images.

		MobileVNet	VGG 19	DenseNet	Incepti-onV3	ResNet 50	SSD	AlexNet	GoogleNet
Mean	Car	0.593	0.486	0.460	0.999	0.739	0.981	0.392	0.697
	Binocular	0.407	0.514	0.540	0.001	0.261	0.019	0.608	0.303
Media-n	Car	0.512	0.429	0.475	1.000	0.818	0.993	0.492	0.824
	Binocular	0.488	0.571	0.525	0.000	0.182	0.007	0.508	0.176
Std Dev	Car	0.223	0.498	0.110	0.001	0.212	0.031	0.199	0.308
	Binocular	0.223	0.498	0.110	0.001	0.212	0.031	0.199	0.308
Majority Vote		tie	tie	car	car	car	binocular	car	car
Conclusion		NA	NA	Likes freedom	Likes freedom	Likes freedom	Focus on big picture	Likes freedom	Likes freedom

Table 3: Results from Experiment 2 (Image C - Figure 2)

		MobileVNet	VGG 19	DenseNet	InceptionV3	ResNet 50	SSD	AlexNet	GoogleNet
Mean	Tree	0.952	0.350	0.882	0.692	0.964	0.233	0.857	0.952
	Hut	0.008	0.498	0.086	0.183	0.001	0.639	0.135	0.048
	Birds	0.039	0.152	0.032	0.125	0.036	0.127	0.009	0.000
Median	Tree	0.949	0.349	0.869	0.739	0.972	0.339	0.975	0.976
	Hut	0.006	0.495	0.083	0.140	0.000	0.607	0.013	0.024
	Birds	0.032	0.150	0.031	0.109	0.027	0.008	0.007	0.000
Std Dev	Tree	0.024	0.067	0.058	0.165	0.037	0.168	0.197	0.057
	Hut	0.008	0.080	0.055	0.132	0.001	0.290	0.195	0.057
	Birds	0.022	0.028	0.015	0.051	0.036	0.161	0.006	0.000
Majority Vote		tree	hut	tree	tree	tree	hut	tree	tree
Conclusion		honest	cheater	honest	honest	honest	cheater	honest	honest

Table 4: Results from Experiment 2 (Image H - Figure 2)

Model	Personality Traits Mimicked	OCEAN Traits
Mobile VNet	logical, intuitive, extrovert, passionate lover, attachment towards mother, fear of taking responsibilities, aware of surroundings, imaginative, practical, cautious, looks at big picture and avoids small details, leadership qualities	C, E, A, N
VGG 19	logical, intuitive, introvert, humorous, likely to cheat, emotional, fear of death, nature lover, practical, cautious, looks at big picture and avoids small details, creative	N
Dense Net	need experience to work on, introvert, humorous, makes undoubted decisions, passionate lover, attachment towards mother, fear of taking responsibilities, keen eye for small detail, imaginative, nature lover, leadership qualities	C, A, N
Incepti-onV3	logical, intuitive, extrovert, passionate lover, emotional, attachment towards mother, fear of taking responsibilities, aware of surroundings, practical, cautious, looks at big picture and avoids small details, creative	C, E, A, N
ResNet 50	need experience to work on, introvert, humorous, passionate lover, emotional, fear of death, nature lover, imaginative, practical, cautious, looks at big picture and avoids small details, creative	C, A, N
SSD	logical, intuitive, introvert, humorous, likely to cheat, emotional, fear of death, nature lover, practical, cautious, looks at big picture and avoids small details, creative	N
AlexNet	logical, intuitive, extrovert, passionate lover, emotional, attachment towards mother, fear of taking responsibilities, nature lover, imaginative, practical, cautious, looks at big picture and avoids small details, leadership qualities	C, E, A, N
Google Net	logical, intuitive, extrovert, makes undoubted decisions, passionate lover, emotional, attachment towards mother, fear of taking responsibilities, nature lover, keen eye for small detail, imaginative, leadership qualities	C, E, A, N

Table 5: Personality Traits Mimicked by the Models

5.3 Results from Experiment 3

In this experiment, the Rorschach test was performed on the image classifiers. Since the test uses abstract ink blot diagrams, training data for each class was handpicked carefully to resemble the ink blots. The ten test images contain options that correspond to 23 different classes. Since each question contains an option of “others”, the sum of probability scores corresponding to all classes not directly present in the image was considered as the probability score of the “other” class. The test attempts to predict how mentally normal a person is in general. Experimental results in Table 6 show all models achieved a score of over 80% normalcy, indicating that none of them makes rash, insensible decisions. With a score of 97%, ResNet50 can be considered to be the ideal model in terms of mimicking human mental stability.

	Test Image Number										Nor-malcy
	1	2	3	4	5	6	7	8	9	10	
Mobile VNet	mask	animal	people	man	moth	boat	witch	ani-mal	fire	spider	85%
VGG 19	bat	animal	people	man	moth	boat	witch	ribs	fire	spider	92%
Dense Net	bat	animal	people	man	moth	boat	witch	ribs	fire	spider	92%
Incept ionV3	mask	animal	people	man	moth	boat	witch	ribs	anat omy	spider	82%
ResNe t50	bat	human	people	man	bat	boat	witch	ribs	fire	spider	97%
SSD	bat	animal	people	man	moth	boat	witch	ribs	fire	spider	92%
AlexN et	mask	animal	people	rug	moth	boat	lady	ribs	fire	spider	82%
Google Net	mask	animal	people	rug	bat	boat	witch	ribs	fire	spider	82%

Table 6: Rorschach's Inkblot Test Results

6. Conclusion

In this work, eight state-of-the-art deep learning image classifiers have been chosen and an empirical analysis is performed to understand if they possess any inherent biases and if they mimic specific human personality traits. In this regard, three psychological object detection experiments (left-brain right-brain test, psychological image test, and Rorschach inkblot) which are usually performed on humans are performed on machines. It was seen that most models work similar to a left-brained person, i.e., they are good at critical thinking, reasoning, logic, math skills, etc. Further, it was seen that the models do not always predict the same class when given images consisting of multiple object classes, the prediction probabilities of each class needn’t necessarily be approximately equal, and, smaller objects are not detected significantly in the presence of larger ones. It is intended to extend this work by curating sophisticated experiments that test more advanced image classifiers like vision transformers and GraphNet on similar lines. Further since this field of study is in its nascent stages, collaborating with psychologists or mental health experts might open up interesting avenues and directions. We also wish to perform a context-sensitive study of images for better classification and object detection.

7. References

Acharjee, S., Abujar, S., Acharjee, S. and Islam, S., 2017. Decision Support System for Online Product Recommendation Service based on Consumer Behavior. *International Journal of Computer Applications*, *975*, p.8887.

Ahmad, S. and Simonovic, S.P., 2006. An intelligent decision support system for management of floods. Water resources management, 20(3), pp.391-410.

Bonini, C.P., 1963. Simulation of Information and Decision Systems in the Firm, Prentiss-Hall. Inc., Englewood Cliff, NJ.

Broca, P., 1865. Sur le siège de la faculté du langage articulé. Bulletins et Mémoires de la Société d'Anthropologie de Paris, 6(1), pp.377-393.

Chen, L., Li, S., Bai, Q., Yang, J., Jiang, S. and Miao, Y., 2021. Review of image classification algorithms based on convolutional neural networks. Remote Sensing, 13(22), p.4712.

Corballis, M.C., 2014. Left brain, right brain: facts and fantasies. PLoS biology, 12(1), p.e1001767.

Del Giudice, V., De Paola, P., Francesca, T., Nijkamp, P.J. and Shapira, A., 2019. Real estate investment choices and decision support systems. *Sustainability*, *11*(11), p.3110.

DENNY, D. and WOLF, R., 1984. Comparison of two personality tests as measures of left-right brain cerebral hemisphere preference and creativity correlates. *The Journal of Creative Behavior*, *18*(2), pp.142-146.

Gomez-Villa, A., Martin, A., Vazquez-Corral, J. and Bertalmío, M., 2019. Convolutional neural networks can be deceived by visual illusions. In Proceedings of the IEEE/CVF conference on computer vision and pattern recognition (pp. 12309-12317).

Gomez-Villa, A., Martín, A., Vazquez-Corral, J., Bertalmío, M. and Malo, J., 2020. Color illusions also deceive CNNs for low-level vision tasks: Analysis and implications. Vision Research, 176, pp.156-174.

Güntürkün, O., Ströckens, F. and Ocklenburg, S., 2020. Brain lateralization: a comparative perspective. Physiological reviews, 100(3), pp.1019-1063.

Haselton, M.G., Nettle, D. and Andrews, P.W., 2015. The evolution of cognitive bias. The handbook of evolutionary psychology, pp.724-746.

He, K., Zhang, X., Ren, S. and Sun, J., 2016. Deep residual learning for image recognition. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 770-778).

He, T., Zhang, Z., Zhang, H., Zhang, Z., Xie, J. and Li, M., 2019. Bag of tricks for image classification with convolutional neural networks. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (pp. 558-567).

Hertz, M.R., 1935. The Rorschach ink-blot test: historical summary. Psychological Bulletin, 32(1), p.33.

Howard, A.G., Zhu, M., Chen, B., Kalenichenko, D., Wang, W., Weyand, T., Andreetto, M. and Adam, H., 2017. Mobilenets: Efficient convolutional neural networks for mobile vision applications. arXiv preprint arXiv:1704.04861.

Huang, G., Liu, Z., Van Der Maaten, L. and Weinberger, K.Q., 2017. Densely connected convolutional networks. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 4700-4708).

Hubbard, K. and Hegarty, P., 2016. Blots and all: A history of the Rorschach ink blot test in Britain. *Journal of the History of the Behavioral Sciences*, *52*(2), pp.146-166.

Hugdahl, K., 2000. Lateralization of cognitive processes in the brain. Acta psychologica, 105(2-3), pp.211-235.

Iaccino, J.F., 2014. Left brain-right brain differences: Inquiries, evidence, and new approaches. Psychology Press.

Ingalhalikar, M., Smith, A., Parker, D., Satterthwaite, T.D., Elliott, M.A., Ruparel, K., Hakonarson, H., Gur, R.E., Gur, R.C. and Verma, R., 2014. Sex differences in the structural connectome of the human brain. *Proceedings of the National Academy of Sciences*, *111*(2), pp.823-828.

Krizhevsky, A., Sutskever, I. and Hinton, G.E., 2012. Imagenet classification with deep convolutional neural networks. Advances in neural information processing systems, 25.

Kukar, M., Vračar, P., Košir, D., Pevec, D. and Bosnić, Z., 2019. AgroDSS: A decision support system for agriculture and farming. *Computers and Electronics in Agriculture*, *161*, pp.260-271.

Li, Zhiyuan, and Sanjeev Arora. "An exponential learning rate schedule for deep learning." arXiv preprint arXiv:1910.07454 (2019).

Liu, W., Anguelov, D., Erhan, D., Szegedy, C., Reed, S., Fu, C.Y. and Berg, A.C., 2016, October. Ssd: Single shot multibox detector. In European conference on computer vision (pp. 21-37). Springer, Cham.

Lonnqvist, B., Bornet, A., Doerig, A. and Herzog, M.H., 2021. A comparative biology approach to DNN modeling of vision: A focus on differences, not similarities. Journal of Vision, 21(10), pp.17-17.

MacNeilage, P.F., Rogers, L.J. and Vallortigara, G., 2009. Origins of the left & right brain. Scientific American, 301(1), pp.60-67.

Mikellides, B., 2012. Colour psychology: The emotional effects of colour perception. In Colour Design (pp. 105-128). Woodhead Publishing.

Nissley, G.E. and DeFreese, E., 2020. Thematic Apperception Test. *The Wiley Encyclopedia of Personality and Individual Differences: Measurement and Assessment*, pp.381-385. Pal, S.K., Pramanik, A., Maiti, J. and Mitra, P., 2021. Deep learning in multi-object detection and tracking: state of the art. Applied Intelligence, 51(9), pp.6400-6429.

Power, T.E., Swartzman, L.C. and Robinson, J.W., 2011. Cognitive-emotional decision making (CEDM): a framework of patient medical decision making. Patient education and counseling, 83(2), pp.163-169.

Sharif, M.I., Khan, M.A., Alhussein, M., Aurangzeb, K. and Raza, M., 2021. A decision support system for multimodal brain tumor classification using deep learning. *Complex & Intelligent Systems*, pp.1-14.

Simonyan, K. and Zisserman, A., 2014. Very deep convolutional networks for large-scale image recognition. arXiv preprint arXiv:1409.1556.

Sutherland, C.A., Rowley, L.E., Amoaku, U.T., Daguzan, E., Kidd-Rossiter, K.A., Maceviciute, U. and Young, A.W., 2015. Personality judgments from everyday images of faces. Frontiers in Psychology, 6, p.1616.

Szegedy, C., Liu, W., Jia, Y., Sermanet, P., Reed, S., Anguelov, D., Erhan, D., Vanhoucke, V. and Rabinovich, A., 2015. Going deeper with convolutions. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 1-9).

Szegedy, C., Vanhoucke, V., Ioffe, S., Shlens, J. and Wojna, Z., 2016. Rethinking the inception architecture for computer vision. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 2818-2826).

Times Now Personality Test: Latest News, Personality Test Videos and Photos - Times Now. https://www.timesnownews.com/topic/personality-test

Vernon, P.E., 1933. THE RORSCHACH INK-BLOT TEST 1. I. British Journal of Medical Psychology, 13(2), pp.90-118.

Wang, X., Hua, X., Xiao, F., Li, Y., Hu, X. and Sun, P., 2018. Multi-object detection in traffic scenes based on improved SSD. Electronics, 7(11), p.302.

Wang, Y., Kitani, K. and Weng, X., 2021, May. Joint object detection and multi-object tracking with graph neural networks. In 2021 IEEE International Conference on Robotics and Automation (ICRA) (pp. 13708-13715). IEEE.

Wong, S.C., Gatt, A., Stamatescu, V. and McDonnell, M.D., 2016, November. Understanding data augmentation for classification: when to warp?. In 2016 international conference on digital image computing: techniques and applications (DICTA) (pp. 1-6). IEEE.

Yahyaoui, A., Jamil, A., Rasheed, J. and Yesiltepe, M., 2019, November. A decision support system for diabetes prediction using machine learning and deep learning techniques. In *2019 1st International Informatics and Software Engineering Conference (UBMYK)* (pp. 1-4). IEEE.

Zeki, S. and Marini, L., 1998. Three cortical stages of colour processing in the human brain. Brain: a journal of neurology, 121(9), pp.1669-1685.