Proposal for a CNET project

1. Summary

I propose to create the nets that, given an input content, will output the content 'seen' by the net in the input.

Further, personalization of these nets will lead to what may be called the 'alter ego' nets. These will simulate perception of an input content and generate the seen content *per person* or group of persons.

The 'alter ego' nets will allow many new applications, like creation of highly personalized content, estimation of human's preferences, etc

In my work on creation of C-nets targeting NLP, I will my explore my experience as a computer scientist and as a poet.

2. Intro

People see the world differently. Even observing the same content. In other words, the 'seen' content differs from the observed, and, the 'seen' is different for different people.

1. the C-nets

1.1.1. definition. examples

I propose to realize my ideas of creating new type neural nets. The nets are trained to perceive an input content and expose an output that may be regarded as the content 'seen by the net' in the input.

One may call such nets the creative nets (C-nets).

One example of C-nets is a net whose input to the net is a scene, and the seen output is the scene with the embedded image of the sought object, that the net has 'seen' in the input (Figure 2,A,B)

Another example of the 'seen' content is textual representation of the input as shown in Figure 2,A,C.

(A) (B)

Figure 1: (A)- observed image (B) - seen image: contains the objects 'seen' in the observed object

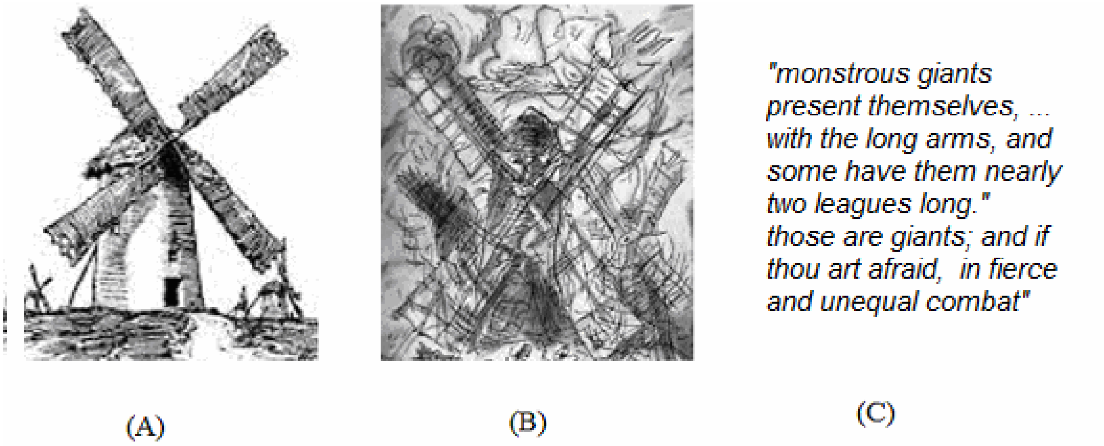

Figure 2: A C-net may simulate perception of Don Quixote in the episode with windmills.
(A): an input image [1]
(B): the perceived image of giants [1]
(C): textual content 'seen' by Don Quixote [2].
The C-nets will be trained on the image pairs like (A,B) and (A,C)

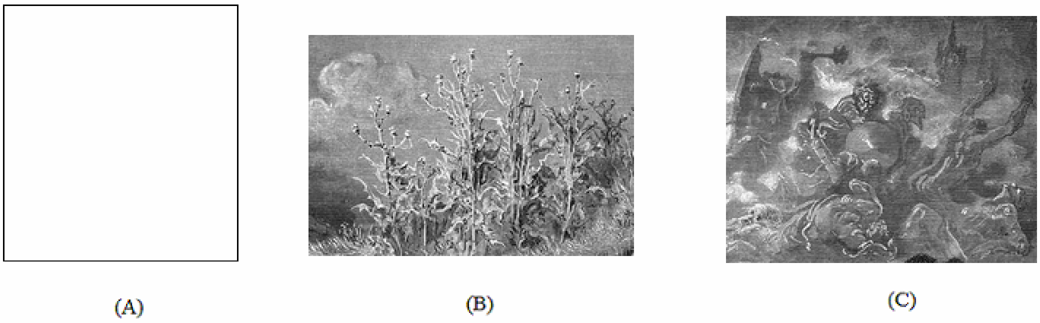

Figure 3: (B) observed [4] (C) seen [3]

The main application of C-nets will be personalized C-nets considered below.

The C-nets are different by

- semantic level. The net with input Figure 1 A and output Figure 1 B is of low level. The net with input Figure 2 A and output 2 B is of high level.
- domain modality. Image domains may be high o low semantic levels. Language domains are of high level.
- observed input may be empty/non empty. Ref to Figure 3 (A) and Table 1.

1.1.2. The C-nets for a natural language content.

The former in main application of C-nets will be personalized C-nets considered below.

1.1.3. Training the C-nets

Training/Testing example is a pair (observed, seen). Basically C-nets will be trained on content pairs as (A,B) or (A,C) of Figure 2.
Visual and textt data are different in that the textual C-net may be trained at very small (eg one) number of examples. For visual data .... to 'blind inpainting'. For textual data the interpretation model is a graph that reflects relations (interactions) between the meanings of words, sounds (probably, also, and letters) comprising esthetically efficient patterns 1.1.5). The model may be tuned on at a single example like the shown in Table 1.

Uchitsya eto ne kopirovat. Dle EE ya mogu buold similar (DPS=>GLO). Ya postroil sililar. Ya mogu tak ze similat to obuchoy kakrtinke. Ponyatie similarity drugoe
VS
EE object is a net. Kogdauchimsya. DElaem TL.(TL et postroenie simil net)
Possibility to learn machine at very small number of examples is property of textual data. objection: ee paintings The data consist of high level semantic objects (words, sounds, letters, their relations and attributes), accumulating human . The number of learning examples is inversially prportional to "semantic height" of data. [When we model them, the work is similar to the use of pretrained net (similar to transfer learning with small target data size). ***ESLI EE object is a net to chto est input of the net?***
duality: zakon net i zakon of person
https://ru.wikipedia.org/wiki/%D0%A1%D1%82%D0%B8%D1%85%D0%BE%D0%B2%D0%B5%D0%B4%D0%B5%D0%BD%D0%B8%D0%B5

indeed we interpret the in terms atributes, links, features.] We do not know whether one may learn visual C-nets at single examples like (B,C) of Figure 3 *without textual interpretation.* [Note: The interpretation models for such images, if exist, differ from DN layer operations models like correlation filters or relus].

1.1.4. Training the textual C-nets

For visual images the net is trained on pairs of (Figure 2 A, B). The trained net should given (Figure 3 A) or (Figure 3 B) yield the image similar to (Figure 3 C).

1.1.5. Training the textual C-nets

<u>Esthetically effective tuples (EE)</u>
We know that sometimes short word tuples invoke unreasonably strong impression to human. For example: (Rain, Steam and Speed) by Turner.

One may explain this effect as follows. These short phrase include a few mutually interacting meanings, the interaction affects the human. The meanings exists at several levels. At low level, interacted are the meanings associated with the letters. At high level, interact the meanings of the words. It is worth to emphasize the low meanings interact with these of high level. The process is similar to perception of visual forms, where detected low level primitives (edges) are interacted with detected high level ones [1], and the whole process is simultaneously bottom-up and top-down.

Humans learn EE at very small number of examples. For example, a new tuple may be output given one EE example (imitation of an input text, Table 1). We may train a C-net in similar fashion using model of EE with free parameters. The model describes relations between the elements of different semanic levels. The parameters may be trained given very small number of "positive" examples. Learning the model parameters resembles learning the NN layers coefficients. [Similarity score for candidates is determined by inexact graph ismorpism.???]

<u>For NLP images</u> the C-net is the net whose input is fixed and is a dictionary of English words. The output is a series of EE triplets. The net will be trained to 'tell' new EE triplets. Collections of high-level literature texts will be used in the training

Level	Learning Example	Evaluation example	Remark	

[1] the 'meaning' here is detection

	Observed	Seen	Observed	Seen		
0	—	Rain Steam speed	—	rain steam speed	repetition	
1	—	rain steam speed	—	gypsum summer faint	imitation	
1	—	the mountains of the far shore swimming in a sunset haze	—	Half-drunk switchman in a red semaphore cap	Creative imitation of specific text	Vladimir Nabokov
2	—	Large text collection	—	Half-drunk switchman in a red semaphore cap	Creative imitation of non-specified text	
3	—	Large text collection	—	Two roads diverged in a wood and I - I took the one less traveled by, and that has made all the difference.	Creation of a new text.	Robert Frost

Table 1: hierarchy of creativity. Machine given learning example, the machine is able to produce evaluation example.

2. *the PC-nets ("alter ego" nets)*

The important subfamily of C-nets will be personalized C-nets (PC) associated with a person or group of persons. The PC nets will be trained so that their perceived content will approximate that of the respective person. In this way, the C-nets will allow tight modeling and classification of the respective human's perception. Thus the IP nets as an technological advance w.r.t. 'static' person's profile.

These may be learned analysis of user feedbacks. (By using transfer learning).

3. Applications of PC nets

Another application of IP nets is that these will may allow creation of highly personalized user content (adversarial, news, summarization).

4. Discussion

3. *why better. cant provebut...*

Why the C-net simulating person (P) is better then profile information associated with (P)?
The C-net exposes approximation to the internal world of P, that should be much more powerful then the profile.
Contextually, the output of C-nets is highly associative/intelligent media .... //why better that plain text?//

2)mozet sami po sebe budut silnee: transfer from intelligence accumulated in the trextsa
1)bolee polnaya simultion of a person. mshiva napevaet kva kva kva.
Mozno observe associativnue relations nu current content

4. *Difference from summarization*

The texts like shown Figure 2 (C) differ from text summarizations of the image (D). The (C) resembles but *is not* summarization of (A). Also, the (B) resembles but *is not* (A).

5. *Difference from deep dream*

5. Org 'How'

6. References

[1] modified images from https://www.google.com/search?as_st=y&tbm=isch&as_q=don+quixote+and+the+windmills&as_epq=&as_oq=&as_eq=&cr=&as_sitesearch=&safe=images&tbs=sur:fmc
[2] http://www.online-literature.com/cervantes/don_quixote/12/
[3] seen: subimage Dore's illustration to Don Chichot https://commons.wikimedia.org/wiki/File:Don_Quixote_2.jpg
[4] observed: https://commons.wikimedia.org/wiki/File:Don_Quixote_4.jpg