applied sciences

Article

Explainable Hopfield Neural Networks Using an Automatic Video-Generation System

Clemente Rubio-Manzano [1,2,*], Alejandra Segura-Navarrete [1], Claudia Martinez-Araneda [3] and Christian Vidal-Castro [1]

1 Department of Information Systems, University of the Bío-Bío, Avda. Collao 1202, Casilla 5-C, Concepción 4051381, Chile; asegura@ubiobio.cl (A.S.-N.); cvidal@ubiobio.cl (C.V.-C.)
2 Department of Mathematics, University of Cádiz, 11003 Cádiz, Spain
3 Computer Science Department, Universidad Catolica de la Santísima Concepción, Concepción 4090541, Chile; cmartinez@ucsc.cl
* Correspondence: clrubio@ubiobio.cl

Abstract: Hopfield Neural Networks (HNNs) are recurrent neural networks used to implement associative memory. They can be applied to pattern recognition, optimization, or image segmentation. However, sometimes it is not easy to provide the users with good explanations about the results obtained with them due to mainly the large number of changes in the state of neurons (and their weights) produced during a problem of machine learning. There are currently limited techniques to visualize, verbalize, or abstract HNNs. This paper outlines how we can construct automatic video-generation systems to explain its execution. This work constitutes a novel approach to obtain explainable artificial intelligence systems in general and HNNs in particular building on the theory of data-to-text systems and software visualization approaches. We present a complete methodology to build these kinds of systems. Software architecture is also designed, implemented, and tested. Technical details about the implementation are also detailed and explained. We apply our approach to creating a complete explainer video about the execution of HNNs on a small recognition problem. Finally, several aspects of the videos generated are evaluated (quality, content, motivation and design/presentation).

Keywords: explainable artificial intelligence; hopfield neural networks; automatic video generation; data-to-text systems; software visualization

Citation: Rubio-Manzano, C.; Segura-Navarrete, A.; Martinez-Araneda, C.; Vidal-Castro, C. Explainable Hopfield Neural Networks Using an Automatic Video-Generation System. *Appl. Sci.* **2021**, *11*, 5771. https://doi.org/10.3390/app11135771

Academic Editor: Federico Divina

Received: 16 May 2021
Accepted: 17 June 2021
Published: 22 June 2021

1. Introduction: AI and Explainability

Presently, transparency is one of the most critical words around the world [1,2]. We perceive it as the quality of easily seeing or understanding the others' actions, implying openness, communication, and accountability [3]. In computer science, transparency has been historically a significant challenge, and it has been addressed in different ways by different disciplines. We can mention here algorithmic transparency, algorithmic accountability [4–8] and, more recently, interpretable/explainable systems [9–16]. The common goal of all the above disciplines is that algorithms' actions must be easily understood by users (expert and non-experts) when we execute them in a particular context.

On the other hand, Artificial Intelligence (AI) is impacting on our everyday lives, improving decision-making and performing intelligent tasks (image and speech recognition, image and speech generation, medical diagnostic systems and more). However, despite AI programs' capabilities for problem-solving, they lack explainability [17]. For example, deep learning algorithms use complex networks and mathematics, which are hard for human users to understand. AI must be accessible for all human users not only for expert ones. Furthermore, it could be a severe problem when an unexpected decision needs to be clarified in critical contexts: medical domain/health-care, judicial systems, banking/financial domain, bioinformatics, automobile industry, marketing, election campaigns, precision agriculture, military expert systems, security systems and education [18]. Therefore, it is a current challenge to understand how and why the machines make decisions. Even the European Union has fostered a framework of regulations for providing the scientific community with guidelines to obtain satisfactory explanations from the execution of algorithms governing decisions' making [19].

On this regard, a current challenge in AI is to achieve explainability in AI [13]. In the literature, several computational methods have been proposed to explain predictions from supervised models [20–22]. On the other hand, there exists works whose objective is to propose visualization and animation techniques. We find different kinds of visualization and animation methods, we can mention here the following:

1. Projections of network outputs to visualize decisions in a K-dimensional space. This approach aims to present a set of images for all training vectors [23].
2. Saliency maps (or heatmaps) to simplify and/or change the visual representation of a picture into something that is easier and more meaningful to analyze [24,25].
3. Colored neural networks to show the changes in the states of neurons [26].
4. Visualization of programs as images using a self-attention mechanism for extracting image features [27].
5. Animations using the Grand Tour technique [28] in which animations are employed to have advantages with respect to classical visualization techniques [29,30].

Despite the facilities shown by the graphical tools, sometimes it is not easy to interpret them, especially when non-expert users are in charge of performing this task. For this reason, Interactive Natural Language Technology for Explainable Artificial Intelligence has been recently proposed [31]: *"the final goal is to make AI self-explaining and thus contribute to translating knowledge into products and services for economic and social benefit, with the support of Explainable AI systems. Moreover, our focus is on the automatic generation of interactive explanations in NL, the preferred modality among humans, with visualization as a complementary modality [...]"*. This paper follows this line of work.

On the other hand, NL is not the only, nor always, most efficient way to have humans communicate with computers. The use of video images in the context of an expert system has shown us that it is not enough to stick pictures to make the communication efficient and attractive [32]. The combination of graphics and text is an effective explanation strategy [33] because the users can focus their attention on the explainer information in the text [34]. At the same time, graphics allows users to form mental models of the information explained in the text. The combination of visual and text information is a useful explanation strategy when it meets four conditions [35]: (1) the text is easy for all users to understand; (2) the visuals are created and evaluated based on the user's comprehension; (3) the visual representations are employed to provide users with good explanations of the texts; and (4) the experience of users with the content is little or no previous. Additionally, video is a ubiquitous data type not only for viewing (reading) but for daily communication and composition (writing) [36].

For all the above considerations, we propose Automatic Video Generation (AVG) as a new interactive NL Technology for Explainable AI whose objective is to automatically generate step-by-step explainer videos to improve our understanding of complex phenomena dealing with a large amount of information, HNNs in our case. The idea is to combine in the same framework explanations in NL with visual representations of the data structures and algorithms employed in a process of machine learning and put them together in a sequence of frames (or slides). The result is a video presentation similar to those created by teachers, experts, or researchers when they need to explain something to an audience. Our approach is strongly context-dependent; hence, we will use this feature to provide users with factual and realistic explanations.

Moreover, storytelling techniques help us to include global commentaries. Additionally, as we have videos (mp4, slides, ppt) as output, the user can classically analyze them: forward, backward, pause, slow, fast, so on. Another significant advantage is that a complete understanding of the models could require several examples. Our approach can generate several videos from different samples quickly using a small number of changes. As this is a first approximation, we have some limitations; we highlight two: (i) the difficulty of generating videos with large neural networks; (ii) the limited degree of sophistication of the sentences generated in natural language. These limitations do not diminish the fact that our proposal is promising, and both limitations will address in future works.

It is important to highlight, although we employ our video-generation method for explaining the execution of a HNN, it could be applied to any problem whose solution involves the implementation of a certain algorithm (classical or machine learning ones) because we use the execution traces to automatically obtain the sequence of frames that will compose the future explainer video. More formally, our research (hypothesis, main objective, methodology, result and evaluation) is explained and formalized in Figures 1 and 2. Additionally, our proposed method is presented and explained in Figures 3 and 4 using block diagrams [37].

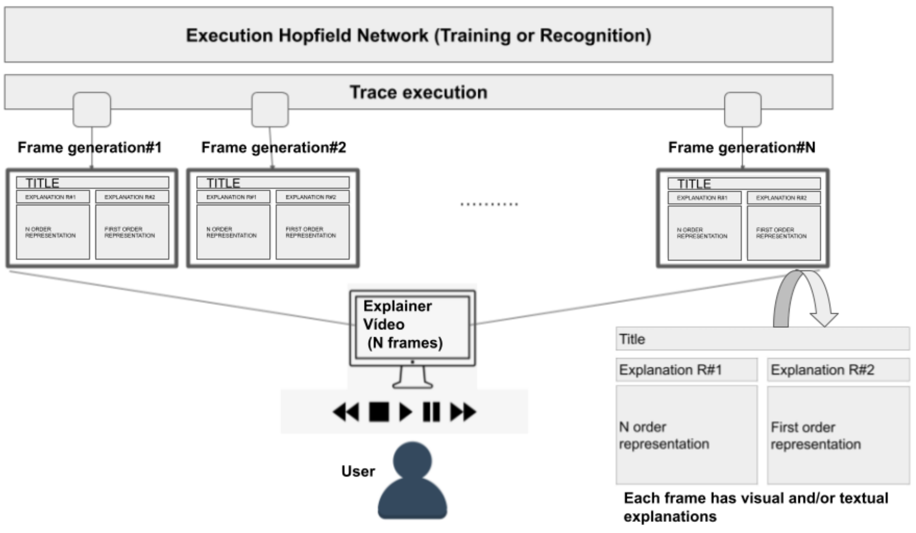
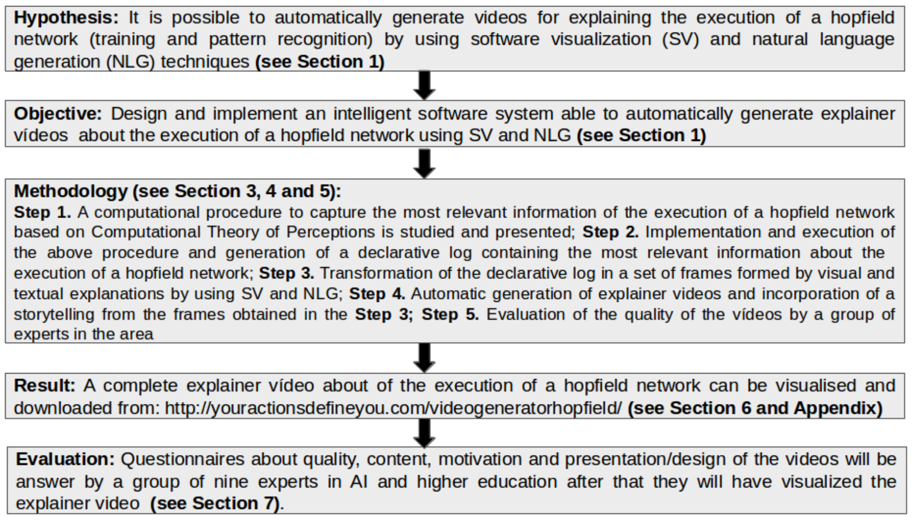


Figure 1. A general framework for transforming the execution of an algorithm in an explainer video using data-to-text and software visualization techniques.



Hypothesis: It is possible to automatically generate videos for explaining the execution of a hopfield network (training and pattern recognition) by using software visualization (SV) and natural language generation (NLG) techniques **(see Section 1)**

Objective: Design and implement an intelligent software system able to automatically generate explainer vídeos about the execution of a hopfield network using SV and NLG **(see Section 1)**

Methodology (see Section 3, 4 and 5):
Step 1. A computational procedure to capture the most relevant information of the execution of a hopfield network based on Computational Theory of Perceptions is studied and presented; **Step 2.** Implementation and execution of the above procedure and generation of a declarative log containing the most relevant information about the execution of a hopfield network; **Step 3.** Transformation of the declarative log in a set of frames formed by visual and textual explanations by using SV and NLG; **Step 4.** Automatic generation of explainer videos and incorporation of a storytelling from the frames obtained in the **Step 3; Step 5.** Evaluation of the quality of the vídeos by a group of experts in the area

Result: A complete explainer vídeo about of the execution of a hopfield network can be visualised and downloaded from: http://youractionsdefineyou.com/videogeneratorhopfield/ **(see Section 6 and Appendix)**

Evaluation: Questionnaires about quality, content, motivation and presentation/design of the videos will be answer by a group of nine experts in AI and higher education after that they will have visualized the explainer video **(see Section 7)**.

Figure 2. Hypothesis, Objective, Methodology, Result and Evaluation.

In the literature, we can find several works about automatic video generation which differ quite a lot from our proposal [38,39]. In [38] a framework for machine learning to generate videos called ArrowGAN is presented and in [39] a multi-modal approach for automatically generating hierarchical tutorials from instructional makeup videos is proposed. The main difference is that in our case videos are automatically generated from execution traces (symbolic information) to obtain a better understanding of machine learning algorithms. In the mentioned approaches videos are employed to generating other videos with different objectives. On the other hand, there is another line of work which aims to automatically generate textual explanations of software projects (code) [40]. This paper is focused on the generation of function descriptions from example software projects and it is applied to study the related problem of semantic parser induction. It employs mainly code as the main resource of communication, no video is generated. To the best of our knowledge, it is the first time that data-to-text based on fuzzy logic and software visualization techniques are employed to automatically generate explainer videos from execution traces.

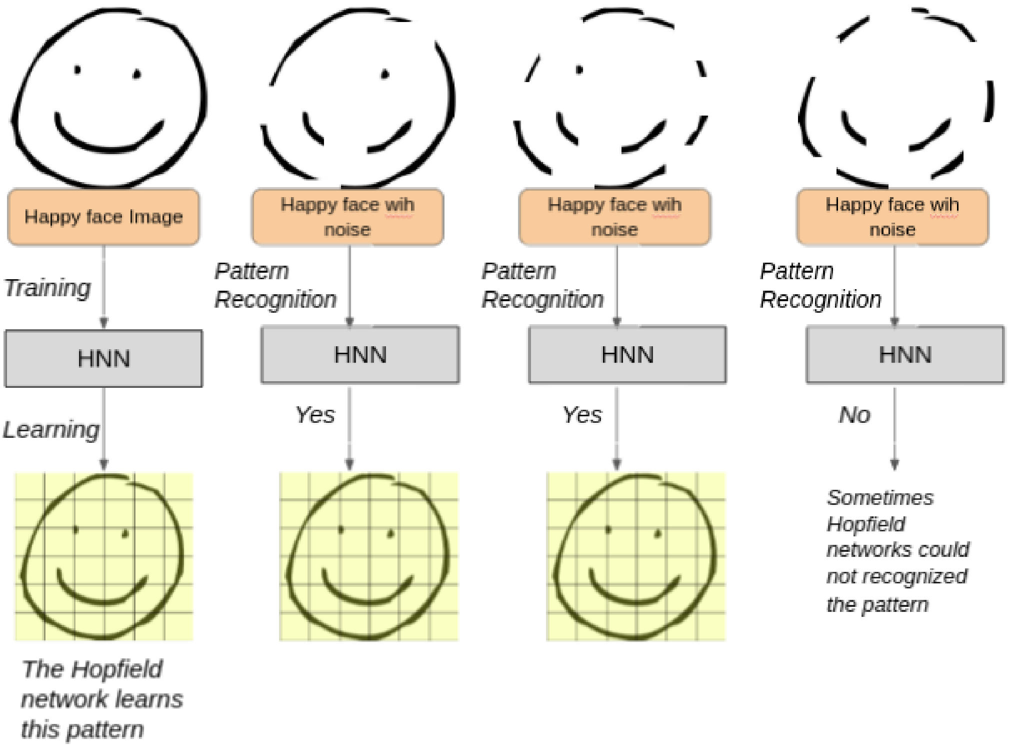


Figure 3. HNN applied to happy-face recognition: training and pattern recognition phases (hypothetical example).

Summarizing, the most relevant contributions of this paper are:

- We propose a new technology to automatically generate explainer videos from HNNs execution traces whose objective is to understand this process better.
- Our approach provides users with visual and textual explanations using data-to-text and software visualization techniques.
- We explain in detail a methodology and propose a software architecture.
- Technical details about the implementation are given and source code is provided on an official web page (youractionsdefineyou.com/videogeneratorhopfield/ (accessed on 19 June 2021)).
- We present an application to generate an explainer video on a real pattern recognition problem.

The rest of the paper is organized as follows. Section 2 introduces several preliminaries concepts about Hopfield neural networks and provides reader with a very brief review of the state of art on the different involved areas. Then, in Section 3 a methodology and a software architecture based on CTP for the automatic generation of explainer videos are presented in detail. In Section 4 are explained the most important task, namely the content determination of frames (how the information will be delivered to the user) and the frames planning (in which order the information will be delivered to the user). Section 5 explains how narration scripts can be used to support the visual and textual explanations during the videos' visualization. In Section 6 technical details about the implementation are given, the main algorithms are presented and explained. Afterwards, Section 7 explains the evaluation carried out on a particular video. Finally, Section 8 provides some concluding remarks and presents the future work with special attention in showing how and why we can use AVG as a powerful teaching-learning resource.

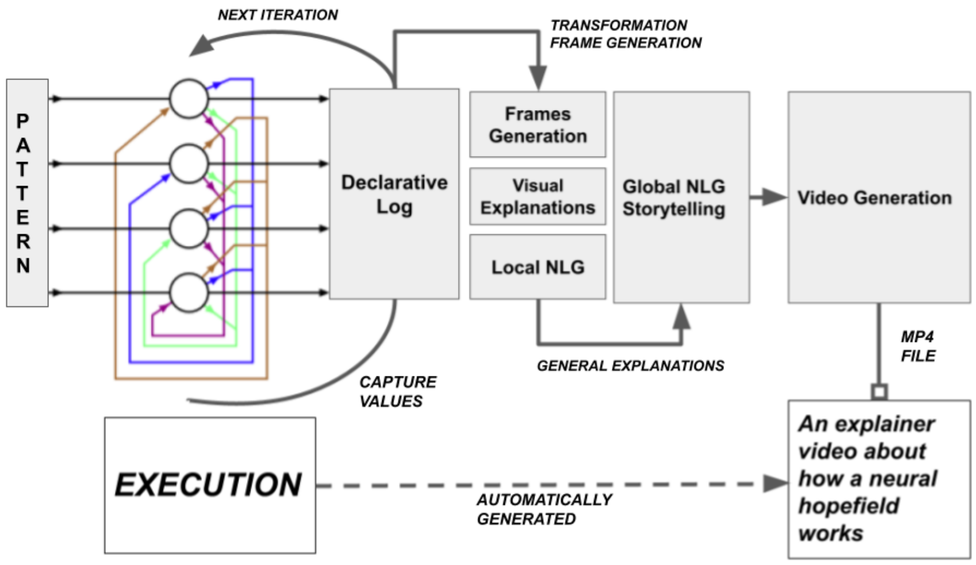


Figure 4. A general methodology for transforming HNN's execution in an explainer video using four phases: trace, frames generation/local NLG, storytelling/global NLG and video generation.

2. Preliminary Concepts

2.1. Hopfield Networks

HNNs are an essential type of recurrent artificial neural networks used to implement associative memory. Their value lies in their ability to pattern recognition, optimization or image segmentation [41]. HNN's origin is the Ising model, a mathematical model employed in the well-known statistical mechanics to explain the ferromagnetism phenomenon. This model use variables to represent magnetic dipole moments of atomic spins. Two states are possible (+1 or -1). This feature is employed in the Hopfield neural networks where each neuron only takes values on two different states (+1 or -1) depending on the input value and if it exceeds their threshold U_i. In particular, Discrete Hopfield neural networks establish a relationship between each pair of neurons (i, j) with $1, 2, \ldots i, j, \ldots N$ [42].

In a particular instant of time, a vector V of N bits is employed to represent the state of the network. Each vector V stores information about which neurons are firing. The interactions between neurons are represented using weights w_{ij} whose values are usually 1 or -1 (We employ this convention in this paper. Other literature employs neurons that take values of 0 and 1.). In this kind of neural networks the Hebb's law of association is employed to learn these interactions, a neuron cannot be connected with itself and the connections between neurons are symmetric. More formally, for a certain state V^s, a weight (i, j) is defined as follows:

$$w_{ij} = (2V_i^s - 1)(2V_j^s - 1)$$

Then $\forall i$ we have that $w_{ii} = 0$ and $\forall i, j$ we have that $w_{ij} = w_{ji}$. Then, the values for each vector are computed as follows:

1. $V_i \rightarrow 1$ if $\sum_j w_{ij} V_j > U_i$

2. $V_i \rightarrow 0$ if $\sum_j w_{ij} V_j < U_i$

The main feature of HNNs is their capability of remembering states stored in the vectors (interaction matrix), i.e., each neuron is going to change until it matches the original state V^s if a new state $V^{s'}$ is subjected to the interaction matrix. A more comprehensive study on the HNNs can be found in [43].

Example 1. *This is a hypothetical example, suppose we want to use an HNN to recognize a happy face in a picture. The first step is to provide the system with a representative example. Once the system learns this pattern, it can recognize any example similar to it. The Figure 3 shows the three options that can occur: (a) training an HNN with an example; (b) recognizing two noise examples and; (c) when the HNN is not capable of identifying an example given.*

2.2. Data-to-Text

Data-to-text systems (D2T) have been gaining relevance in AI during recent years. These systems are working in real-life applications (Galiweather [44] or Metoffice [45]) obtaining very promising results [46,47]. We are referring here to D2T as a discipline that includes three areas: Theory of Computational Perceptions (CTP), Linguistic Descriptions of Complex Phenomena (LDCP) and Natural Language Generation (NLG). CTP [48,49] is based on the way human beings use NL to reason, make decisions and communicate their experience in a context. The main feature of CTP approaches is that they can work in environments in which there exist imprecision, uncertainty and partial truth. CTP aims to develop intelligent computational systems capable of computing and reasoning with imprecise descriptions in NL in a similar way as humans do. On the other hand, the field of the Linguistic Description of Complex Phenomena (LDCP) theory [50] aims to automatically generate linguistic reports using a conceptual framework based on CTP for the automatic generation of reports using natural language about a well-defined phenomenon. There are several real-life applications based on the LDCP framework, such as [51–53]. Finally, NLG systems [45] aims to transform databases in reports in natural language. The main difference between classical NLG systems and fuzzy approaches is the management and treatment of imprecision/vagueness. The first one allows us to obtain more elaborated reports in NL and the treatment of vagueness need to be Incorporated in the system. In the second one, the treatment of vagueness is in the core, but the style of the linguistic reports is less rich. Both proposals have been evolving in a parallel way and they are complementary rather than competitive.

2.3. Film Generation, Software Visualization and 2D Programming

Like Natural Language Generation, Film Generation aims to better comprehend a particular domain by automatically generating a film. Ethnologists [54] have identified at least two levels of film comprehension. The first level is concerned with understanding single images (determining what is in the picture, differentiation of objects in them, confusion between images and reality). The second one concerns the comprehension of the storyline. Film generation is related to software visualization (SV). SV is a visual computational technique usually employed in analysis and development of programs, software maintenance, reverse engineering, and collaborative development. SV aims to propose graphical representations (static or dynamic) of data structures, algorithms or programs (code) to obtain a better understanding of them when they are dealing with large amount of information. The main challenge of this discipline is to obtain good graphical representations using visual metaphors to explain different aspect of the software development. To implement these graphical representations, we need to use a library for 2D programming. We are going to use the Java Foundation Classes (JFC) which is a standard API for providing programmers with a graphical user interface (GUI) for Java programs. In particular, we employ all the functionalities of the well-known Abstract Window Toolkit (AWT). AWT is a library which allows programmers to graphically draw simple or complex representations in two dimensions using a set of graphical primitives. In this paper, we employ the following primitives (Please note that other primitives could be used to enrich the graphical representations of our approach, it will depend on the available time for designing and implementation of each AVG system):

- `drawString(string,x,y);` allows us to draw on the screen a string in the position (x, y).
- `drawRect(x,y,width,height);` allows us to draw on the screen a rectangle in the position (x, y) with a particular width and height.
- `fillRect(x,y,width,height);` allows us to draw on the screen a filled rectangle in the position (x, y) with a particular width and height. The color of the fill can be modified by the following function.
- `setColor(Color);` allows us to change the current color employed in the current privative. Colors can be changed anytime, anywhere.
- `setFont(Font);` allows us to change the font employed in the current `drawString` privative. Fonts can be changed anytime, anywhere.

A useful feature of our approach is that more complex graphical representations can be created by combining these simple primitives (see Section 6 and the function called `drawPattern` for more detail). Additionally, we will use the graphviz software which is as the authors said: *"open-source graph visualization software. Graph visualization is a way of representing structural information as diagrams of abstract graphs and networks. It has important applications in networking, bioinformatics, software engineering, database and web design, machine learning, and in visual interfaces for other technical domains."* [55].

3. Automatic Explainer Video Generation

In this section, we introduce the philosophy behind automatic explainer video generation. Additionally, we focus on two crucial aspects: (i) how to design an AVG system based on CTP and software visualization and; (ii) how should be implemented these kinds of systems. We propose a methodology to answer to item (i), and software architecture to answer the item (ii).

3.1. The Design of Video-Generation System Based on CTP

Our design is based on two crucial aspects: experience and human-computer interface which were pointed out in the work [56].

1. Experience. In this context, the experience [57] can be defined as the descriptions employed to explain a complex phenomenon which is formed by two levels of reality. The first level (also called first-order phenomena) is directly related to the environment. The second one (also called second-order phenomena) is formed by the meanings and wordings of the perceptions of the first-order one. The desired result is to obtain AI systems (personal assistants) capable of helping us to automatically create our personal experience. For defining the experience in HNNs, we will employ first-order representations related directly to the domain, i.e., algorithms and data structures used to implement them. And the second-order ones formed by graphical models and abstractions of the first ones to facilitate their understanding. Hence, our idea is to obtain an artificial teacher assistant capable of explaining the execution of HNNs and provide users with good explanations about the process of training and patten recognition.
2. Human-Computer Interface. We must consider three fundamental items when our objective is to communicate meaning with NL using an intelligent and automatic personal assistant system [57]. First, we need to establish a relation between the users and the AI system. Second, we must define the domain of language between them. Third, a mode of expression must be defined (informative, persuasive, didactical, etc.). Additional features could be taken into account. For example, the use of graphics and sounds could be considered to complement explanations in NL or; the capability of expressing emotions with the tone of voice and the facial expressions of an avatar. Our human-computer interface is an explainer video formed by the following elements: (i)

concepts related to HNNs (matrix for representing the activation of neurons, graph for illustrating the connections between neurons, so on); (ii) a student-teacher relationship; (iii) the mode of expression will be a set of frames formed by explanations in NL and complementing it with the use of graphics. This video presentation aims to be informative and didactic.

The design and the implementation of our human-computer interface will perform using a combination of two computational disciplines, namely: data-to-text systems (CTP, LDCP and NLG) and software visualization using 2D programming. Each one of them is dedicated to performing a particular task. In the following, we clarify the tasks performed by each component.

- **Data-to-text (CTP + LDCP + NLG).** First, we employ CTP as a tool for knowledge representation capable of capturing a human expert's perceptions, in this case, a teacher whose objective is to teach and explain HNNs. As we mentioned, our philosophy is creating explainer videos in a similar way than a human teacher creates them using a set of slides formed by pictures and explanations in NL. Secondly, we use LDCP to define two kinds of computational perceptions, but now we will call them visual representations (first- and second-order representations). The first ones will represent internal structures as variables or data structures. The second-order ones represent more abstract visual representations, and we define them from the first-order ones in a similar way than we calculate n order perceptions from first-order ones in LDCP. Finally, we employ NLG to generate pieces of sentences in natural language. We adapt the techniques investigated in this discipline to create sentences in NL from visual representations. Additionally, we establish the content determination and planning phases to work with pictures and sentences.
- **Film Generation using software visualization and 2D Programming.** We start with a representation of shots frames. The main challenge is finding good graphical representations using visual metaphors and NL explanations. We aim to combine both communication mechanisms by creating visual and textual structures. This technique is employed to generate 2D graphical representations using 2D programming. We will transform the symbolic terms captured in the execution trace in a set of 2D visual drawings which will incorporate into the process of video generation.

3.2. A Software Architecture to Build AVG Based On CTP

We propose a software architecture based on the previously presented methodology. Four modules form the architecture: declarative log, frame generation/local NLG, global NLG/storytelling and video generation.

1. **Declarative log.** As is well known, traceability of neural network training is a very complex task because it employs so many neurons, connections and weights what it implies so many iterations. This module aims to capture the values of the data structures generated during an HNN execution and generate a declarative log from them that the next modules can use. Information about neurons, connections and weights will store using an ontology [58]. The reason for using a declarative log is double: first, we have a formal knowledge representation of the execution process, and second, it can help users to interpret the captured in the next phases, for example, when the system perform the content determination and planning phases. More formally, we define a symbolic video frame using a fact with the following form `frame(id, action,[data_structures])`; where `action` indicates the action performed in that instant of time identified by `id` and `[data_structures]` are the values of each data structure associate to a particular action. Therefore, the complete execution of an HNN is as follows:

```
frame(1,action_1,[data_a1]).
frame(2,action_2,[data_a2]).
...
```

```
frame(n,action_n,[data_an]).
```

We explain the transformation of this set of symbolic frames to visual ones in Section 4. Additionally, we describe technical details about the implementation in Section 6.

2. **Frames generation and local NLG.** This paper adopts the classical architecture of NLG systems, but we adapt it for automatic frames generation. An essential feature of our approach is that the system generates NL descriptions from internal representations of variables and data structures. It consists of the following steps: the content determination subsystem computes an explainer model of the result of the execution of HNNs. It employs two kinds of visual explanations: first-order representations for visually representing internal variables and data structures; and second-order ones for more abstract models of them. The system creates second-order ones from the first ones. It is fed to a planning system and identifies which parts of these representations must be explained and in what order.
3. **Global NLG—Storytelling.** This module implements the "voice" of the expert or teacher. The aim is to find the perfect combination of narrative and data. The idea is to employ storytelling as it is used in other disciplines as training and education contexts: to motivate and to illustrate. The standard approach to incorporating storytelling into a computational intelligent system are the well-known scripts. We employ a script based on classical presentations, i.e., we define the first frame with a little explanation about the presentation's context. Next, we put an index about the main items that we will explain during the video. Finally, we define a slide for the last remarks. On the other hand, we need to be capable of defining extra explainer slides on groups of local slides. This process can be seen as a global NLG system which takes as input a set frames (texts and visual explanations) and automatically create slides. Each extra slide is putting before them with an explanation about that group of slides. For example, an extra explainer slide for the training phase is the one created from the explanations of a group of slides generated to explain the training process. The same process could be performed for the pattern recognition phase (see Section 5).
4. **Video Generation.** This module aims to generate a video from a set of frames automatically. In this work, we do not address the automatic selection of the type of letters, colors, discourse, etc.

4. Frames Generation and Local NLG

4.1. Content Determination of Frames

In this section, we will explain the content determination of the frames. Content determination identifies the information we want to communicate through the generated explainer video. Therefore, video frames will be a set of explanations in NL and graphical representations. The central concept here is the neural network. To define a neural network data type, we need to define two types of data: a set of neurons and a set of connections between them. Therefore, each frame's content provides users with information about these types' values. We employ here the UML language for defining the data types and their relations.

In particular, Figure 5 shows the different data types and their relations (Neuron, Connection, Neural Networks). Each of them contains fields (attributes or properties), and code in procedures (often known as methods). We represent and store the information generated during the HNN execution in a declarative log (ontology). We define each fact "frame" as a tuple of four elements: a unique identifier that indicates the order of execution; a label action gives us information about which operation was performed (initialization, neurons, connections, iteration, so on); information about the values of the data structures employed by the HNN and in some cases a set of messages. Different kind of messages can be generated from each frame. For example, the execution of an HNN can generate a declarative log of this kind (we omit some details for simplicity):

```
frame(1,neurons,[N0,N1,N2,N3,...,N6,N7,N8,N9],
                 "The neurons defined for the
                 Hopfield network are:").

frame(2,connections,[(N0,N1,0.0),...,
                                 (N9,N9,0.0)],
                "The connections between
                neurons have been
                established:").

frame(3,pattern([1.0,1.0,1.0,-1.0,-1.0,
                -1.0,-1.0,-1.0,-1.0,-1.0],
                "The input pattern is:")).

frame(4,iteration#1,[w(0,1,1.0)],
            "Neuron N0 is enable.
             The neuron N0 is
             connected with
             neuron N1 and N2").

frame(5,iteration#2,[w(0,2,1.0),
                    (w(1,2,1.0)]).
...
frame(13,iteration#9,[(w,0,9,-1.0), ...
                      (w,8,9,1.0)]).
```

Additionally, we employ graphical representations to provide users with useful and visual explanations about the changes produced in the internal states of an HNN during the execution. These representations will be essential in the next phases to obtain simple, informative and intuitive messages about the neurons, connections, neural network and training/pattern recognition tasks. We define two types of visual explanations, namely: first-order representations to visually explain graphical low-level data structures, variables or values involved in the process of execution; and the second-order ones to model more abstract figures. An essential feature of our approach is that the system generates second-order ones from the first ones automatically in a similar way than computational perceptions are created in the LDCP paradigm. The Figure 6 shows both types of representations. We represent neurons, connections, and weights as a matrix of weights on the right side. On the left side, a classical graph represents the neurons' activation, vertices are the neurons, edges are the connections, and the labels are the weights. Each visual explanations will be associated with a set of messages explaining the pieces of information shown in them.

4.2. Frames Planning

In essence, a plan is a sequence of actions that bring a system from an initial state to a goal state. A common approach to AI planning involves defining a series of atomic steps to achieve a communication goal. When the communication goal is to explain an execution of HNNs, then the plan is partially guided by the implementation and the execution trace. This last one provides the planner with a sequence ordered of frames. In our case, the steps needed to explain the HNN execution it is as follows (see Figure 7):

1. **Initialization**. The system (using content determination resources) visually shows and textually explain the neurons, connections, neural networks weights and input pattern.
2. **Training**. The system (using content determination resources) visually show and textually explain the training procedure step by step. The system can employ a set of frames to provide user with useful information about it.
3. **Pattern Recognition**. The system (using content determination resources) visually show and textually explain the pattern recognition procedure step by step. The system can employ a set of frames to provide user with useful information about it.

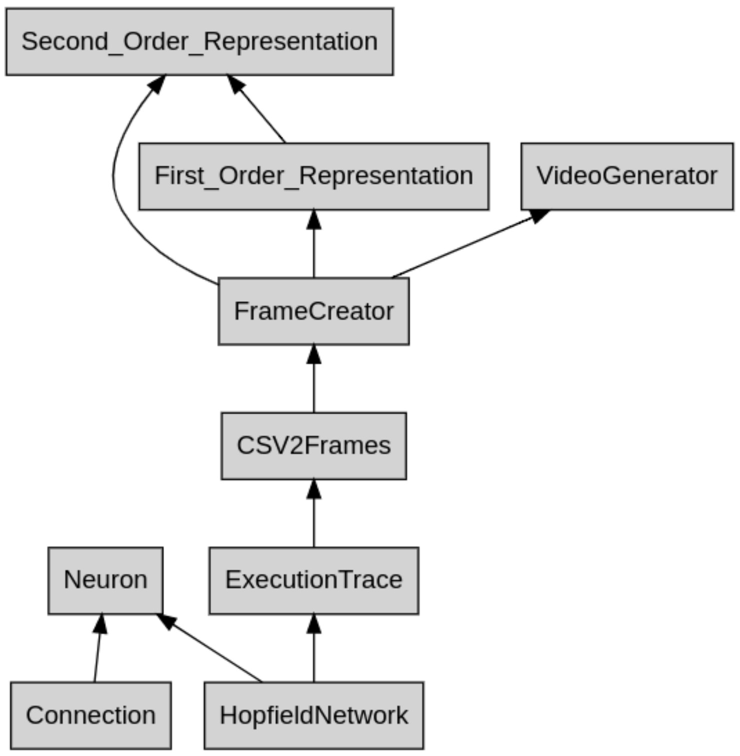
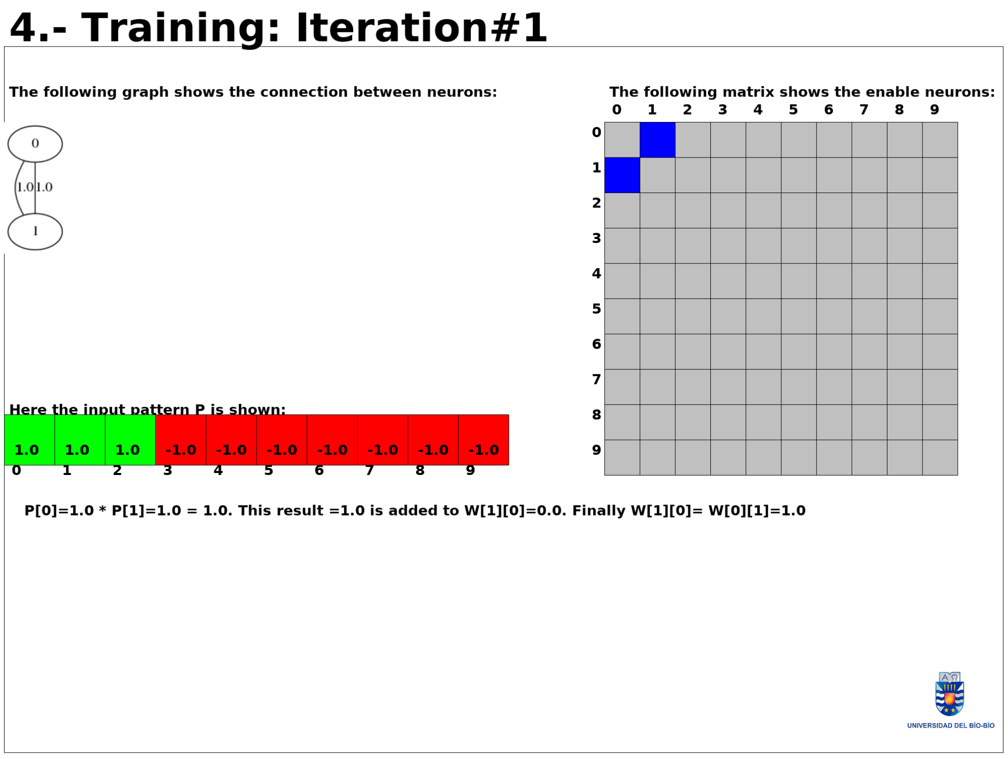


Figure 5. Diagram about the relations between the different concepts and their relations. Arrows represent relation between classes.



4.- Training: Iteration#1

The following graph shows the connection between neurons:

The following matrix shows the enable neurons:

Here the input pattern P is shown:

1.0	1.0	1.0	-1.0	-1.0	-1.0	-1.0	-1.0	-1.0	-1.0
0	1	2	3	4	5	6	7	8	9

P[0]=1.0 * P[1]=1.0 = 1.0. This result =1.0 is added to W[1][0]=0.0. Finally W[1][0]= W[0][1]=1.0

Figure 6. Frame automatically generated in which textual explanations are combined with graphical representations (first-order and second-order).

Additionally, as the system using a particular plan to simulate a teacher's presentation, it incorporates some extra frames before explaining the training and pattern recognition tasks. We add three slides more to the original sequence: Initial presentation (title, contact, date, so on); an index with little explanations of each item and; a slide (or several) for conclusions.

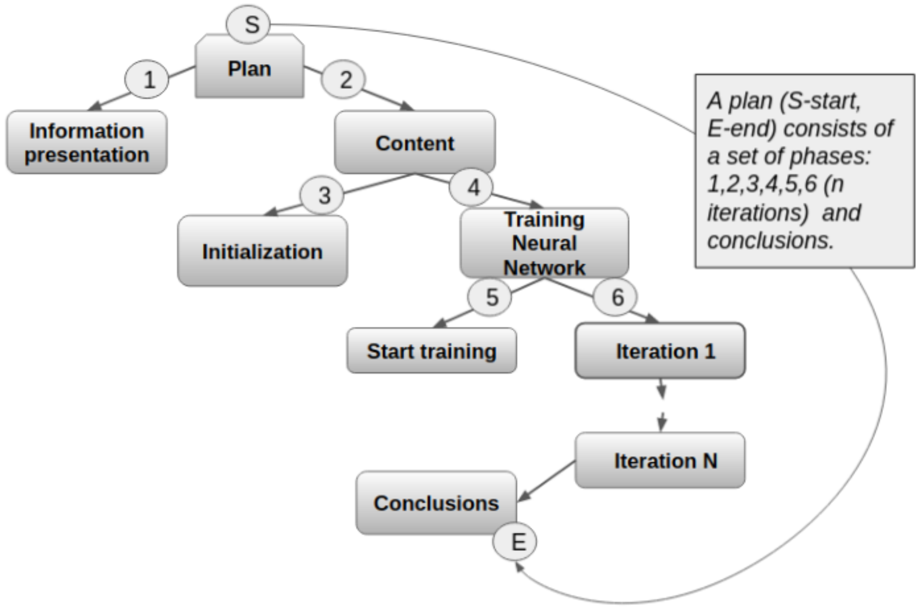


Figure 7. Planning for the video presentation of HNNs execution (training).

5. Global NLG—Storytelling and Narrative

This section explains how to employ a narration script for supporting the visual and textual explanations during the videos' visualization. A script is a standard approach to incorporating storytelling into a computer system. In [59] some recommendations are given to obtain a good narrative in the presentations:

- **Good examples.** The use of illustrative and specific examples to show your ideas and explain the scope of the talk in the time available.
- **Detective story.** The narrative style must follow a detective story structure. The speaker begins by presenting the specific problem and then he/she describes the search for a solution.

We aim to find the perfect combination of data and narrative. We use a narrative style similar to the one used in training and education contexts to illustrate and to motivate. Our storytelling module is the "voice" of the artificial expert or teacher. In our case, we employ a template-based approach. We focus on the logical causal progression of representations, as it is mentioned in [60]: *"Causality refers to the notion that there is a relationship between temporally ordered events. One event changes the story world in a particular way that enables future events to occur"*. In our case, a story expresses essentially how and why an HNN changes. The HNN storytelling consists of explaining the current changes produced during the execution, taking into account the past ones and establishing some relations (possibly casual) between them.

In its current version, our template script is straightforward. We establish a video script in two steps. First, we define each group of the slides from the items defined in the index slide. Second, we create an explainer slide for and from a group of slides. This special slide aims to summarize all the explanations generated to explain a complex process (for example, training or pattern recognition). We use here a first prototype of an avatar for simulating a human teacher. The storytelling is optional; the users can choose if they want to extra explanations about some slides. These slides could remove by the user or using as support notes. Figure 8 summarizes this process.

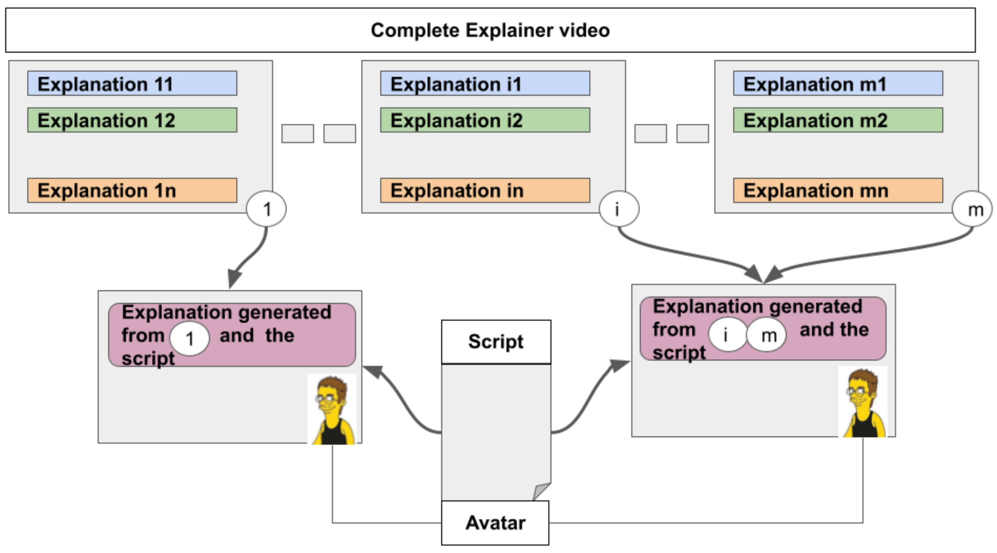


Figure 8. Storytelling pipeline.

6. Technical Details about the Implementation

In this section, technical details about the implementation are given. The first implementation has been developed in Java 11. Several explainer videos and the source code can be watched and downloaded from the official web page of this paper (youractionsdefineyou.com/videogeneratorhopfield/, accessed on 19 June 2021). The complete implementation is summarized using UML diagrams (see Figure 9).

The implementation is formed by two packages: Hopfield neuronal network and video generation. The first one implements all the functionality required to execute the HNN (training and pattern recognition). The execution of an HNN consists of performing three steps: (i) to receive as input an array of 1 and −1 (the pattern); (ii) to train the HNN for recognizing this pattern; and (iii) to recognize any array received as input. This package has a class called `ExecutionTrace` for storing all the information about the HNN execution. The result is a structured file containing a sequence of rows with information about the execution. Each row is formed by an identifier (`ID`) that indicates the order in which a row must be generated; a label action (`LABEL_ACTION`) gives us information about which operation was performed (initialization, creation neurons, creation connections, creation neural networks, weights update, so on). After that, a set of values `D1; ...; DN` is provided with information about the results obtained after the execution of the action indicated previously. Finally, a set of messages about each D_i could be also indicated to support the process of the video generation.

```
ID0;Label_Action;D1; D2;...;DN; M1;...;MN;
ID1;Label_Action;D1; D2;...;DM; M1;...;MM;
...
ID1;Label_Action;D1; D2;...;DL; M1;...;ML;
```

The video-generation package implements the functionality needed to generate an explainer video from the information created by the execution trace package. The main classes are CSV2Frame, FrameCreator, First-Order Representation, Second-Order Representation and VideoGenerator. The main Algorithm A1 receives as input an execution trace (list of rows) and returns a video in mp4. It reads one by one each line and creates a slide using the information associated with each row calling to the function `Create-Slide(row[])` (Algorithm A1-line 5 in Appendix A).

The function `Create-Slide(row[])` creates an empty slide, and it will call a sequence of functions whose objective is to create, where appropriate, a frame with visual and textual explanations. The mentioned functions are:

- `first-order-representations`
- `second-order-representations`
- `explanations-NL-for`

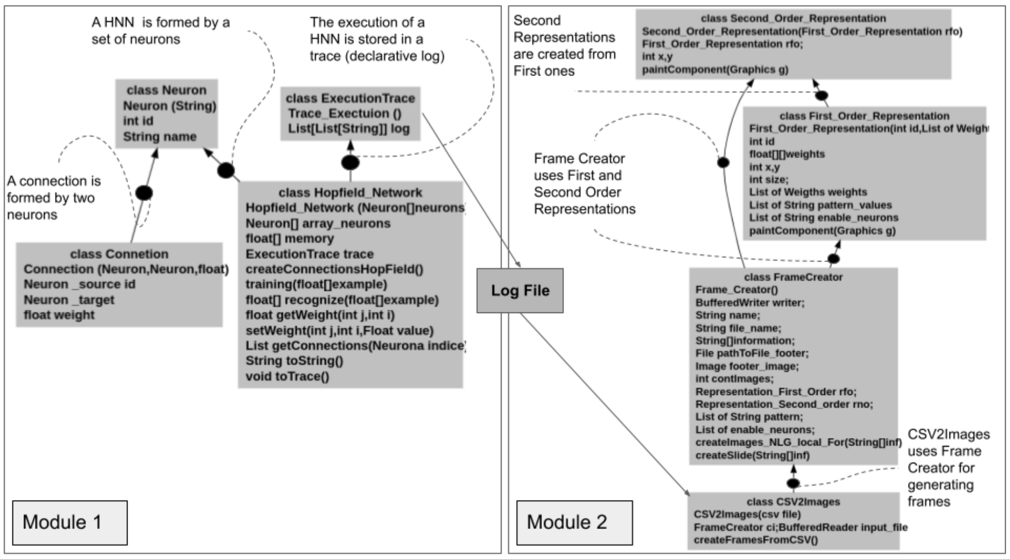
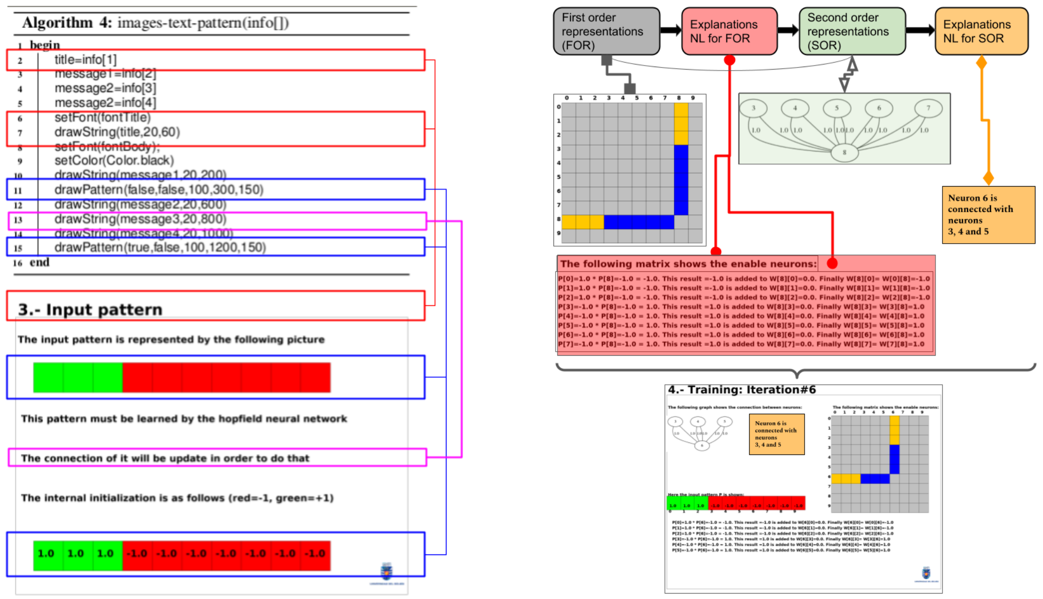

A HNN is formed by a set of neurons
The execution of a HNN is stored in a trace (declarative log)
class Neuron
Neuron (String)
int id
String name
class ExecutionTrace
Trace_Exectuion ()
List[List[String]] log
A connection is formed by two neurons
class Hopfield_Network
Hopfield_Network (Neuron[]neurons
Neuron[] array_neurons
float[] memory
ExecutionTrace trace
createConnectionsHopField()
training(float[]example)
float[] recognize(float[]example)
float getWeight(int j,int i)
setWeight(int j,int i,Float value)
List getConnections(Neurona indice)
String toString()
void toTrace()
class Connetion
Connection (Neuron,Neuron,float)
Neuron _source id
Neuron _target
float weight
Log File
Module 1
Second Representations are created from First ones
class Second_Order_Representation
Second_Order_Representation(First_Order_Representation rfo)
First_Order_Representation rfo;
int x,y
paintComponent(Graphics g)
class First_Order_Representation
First_Order_Representation(int id,List of Weigh
int id
float[][]weights
int x,y
int size;
List of Weigths weights
List of String pattern_values
List of String enable_neurons
paintComponent(Graphics g)
Frame Creator uses First and Second Order Representations
class FrameCreator
Frame_Creator()
BufferedWriter writer;
String name;
String file_name;
String[]information;
File pathToFile_footer;
Image footer_image;
int contImages;
Representation_First_Order rfo;
Representation_Second_order rno;
List of String pattern;
List of enable_neurons;
createImages_NLG_local_For(String[]inf)
createSlide(String[]inf)
CSV2Images uses Frame Creator for generating frames
class CSV2Images
CSV2Images(csv file)
FrameCreator ci;BufferedReader input_file
createFramesFromCSV()
Module 2

Figure 9. Modules and their diagrams of classes (main elements, attributes and methods for each class).

The first and second functions generate visual representations from the information received as parameters. The third function generates textual explanations from the visual ones. It is important to mention that sometimes we do not employ this function in the current implementation, but it will be an essential resource of explainability in the future. The Algorithm A2 shows a pseudocode to create a new frame and Algorithm A3 shows how we implement the function `first-order-representations`. As we can observe this function creates a frame depending on the action received as parameter. For example, we explain here two of them: a simple one (`images-text(pattern,info)`); and a more complex one (`images-text(iteration,info)`) which is called complete cycle of representation (see Appendix B for more details).

The function `images-text(pattern,info)` generates an explainer frame about the input pattern from an array of information `info[]` that contains the explanations generated to explain the data structures employed to implement this element. This function uses 2D atomic primitives to show the messages and create a graphical representation of the input pattern. The Figure 10 graphically explains this process. Sometimes atomic primitives are not enough to implement a particular graphical representation, and we need to create additional functions. For example, we define the function `drawPattern` that allows us to draw a visual matrix with different options: matrix alone, showing indexes or showing values. The possibilities will indicate in the parameters. If wn=true, then it creates a visual matrix showing their values; if wi=true, then it creates a graphical matrix with indexes (see Figure 10).

For the generation of a group of frames whose objective is to explain a particular phase's execution, the system must perform a complete cycle of representation, i.e., for each frame, it generates a first-order model from the information provided by the neurons, connections and the matrix of weights in a particular instant of time during the execution; and textual explanations about that. If the new representation needs to be clarified (it is a designer's decision), the system can generate additional descriptions. Next, the system generates a second-order model from the first one. It aims to provide users with an alternative and more friendly representation of the first one (the system can also generate explanations for this representation). The Figure 10 shows how this cycle is performed, and the concrete details about the implementation can be consulted in the source code on the official web page.



```
Algorithm 4: images-text-pattern(info[])
begin
    title=info[1]
    message1=info[2]
    message2=info[3]
    message2=info[4]
    setFont(fontTitle)
    drawString(title,20,60)
    setFont(fontBody);
    setColor(Color.black)
    drawString(message1,20,200)
    drawPattern(false,false,100,300,150)
    drawString(message2,20,600)
    drawString(message3,20,800)
    drawString(message4,20,1000)
    drawPattern(true,false,100,1200,150)
end
```

3.- Input pattern

The input pattern is represented by the following picture

This pattern must be learned by the hopfield neural network

The connection of it will be update in order to do that

The internal initialization is as follows (red=-1, green=+1)

1.0	1.0	1.0	-1.0	-1.0	-1.0	-1.0	-1.0	-1.0	-1.0

The following matrix shows the enable neurons:

```
P[0]=1.0 * P[8]=-1.0 = -1.0. This result =-1.0 is added to W[8][0]=0.0. Finally W[8][0]= W[0][8]=-1.0
P[1]=1.0 * P[8]=-1.0 = -1.0. This result =-1.0 is added to W[8][1]=0.0. Finally W[8][1]= W[1][8]=-1.0
P[2]=1.0 * P[8]=-1.0 = -1.0. This result =-1.0 is added to W[8][2]=0.0. Finally W[8][2]= W[2][8]=-1.0
P[3]=-1.0 * P[8]=-1.0 = 1.0. This result =1.0 is added to W[8][3]=0.0. Finally W[8][3]= W[3][8]=1.0
P[4]=-1.0 * P[8]=-1.0 = 1.0. This result =1.0 is added to W[8][4]=0.0. Finally W[8][4]= W[4][8]=1.0
P[5]=-1.0 * P[8]=-1.0 = 1.0. This result =1.0 is added to W[8][5]=0.0. Finally W[8][5]= W[5][8]=1.0
P[6]=-1.0 * P[8]=-1.0 = 1.0. This result =1.0 is added to W[8][6]=0.0. Finally W[8][6]= W[6][8]=1.0
P[7]=-1.0 * P[8]=-1.0 = 1.0. This result =1.0 is added to W[8][7]=0.0. Finally W[8][7]= W[7][8]=1.0
```

First order representations (FOR)
Explanations NL for FOR
Second order representations (SOR)
Explanations NL for SOR
Neuron 6 is connected with neurons 3, 4 and 5
4.- Training: Iteration#6
The following graph shows the connection between neurons:
The following matrix shows the enable neurons:
Neuron 6 is connected with neurons 3, 4 and 5

Figure 10. (**Left**) Frame generated using Algorithms A3–A5; (**Right**) A complete cycle of representation for a frame extracted from a group of them whose objective is to explain the training phase.

7. Evaluation

In this section, we evaluate different aspects of the explainer videos created using our method. In order to do this, we are going to employ a standard evaluation instrument called LORI (Learning Object Review Instrument) [61]. This instrument allows us to evaluate relevant aspects of the videos by applying multiple-choice questionnaires about them: video quality, content quality, motivation and design/presentation. The questionnaires will be answer by a group of nine experts in AI and higher education after that they will have visualized the explainer video.

We have designed and created four kinds of questionnaires formed by several questions related to the mentioned aspects (see Tables 1–4). The possible answers go from Strongly Agree (5) to Strongly Disagree (1).

Table 1. Questions and answers to evaluate Video quality.

Questions	Answers					Average
	1	2	3	4	5	
$Q_{1.1}$	0	1	0	4	4	4.2
$Q_{1.2}$	0	0	2	4	3	4.1
$Q_{1.3}$	0	1	2	4	1	3.2
$Q_{1.4}$	0	0	1	4	4	4.3

Table 2. Content.

Quality Content	Answers						Avg
	1	2	3	4	5	None	
$Q_{2.1}$	0	1	2	4	1	0	3.2
$Q_{2.2}$	0	0	2	5	1	1	3.4
$Q_{2.3}$	0	0	2	5	1	1	3.2
$Q_{2.4}$	0	1	1	2	4	0	3.6
$Q_{2.5}$	0	1	0	2	3	2	2.5

Table 3. Motivation.

Motivation	**Answers**						**Average**
	1	**2**	**3**	**4**	**5**	**None**	
$Q_{3.1}$	0	0	5	3	0	1	3.0
$Q_{3.2}$	0	4	1	1	2	1	2.2
$Q_{3.3}$	0	1	0	4	4	0	4.2

Table 4. Design and presentation.

Design and Presentation	**Answers**						**Average**
	1	**2**	**3**	**4**	**5**	**None**	
$Q_{4.1}$	0	1	0	6	2	0	3.8
$Q_{4.2}$	0	0	1	4	2	2	3.2
$Q_{4.3}$	5	0	0	0	1	3	1.1
$Q_{4.4}$	0	1	1	4	2	1	3.4
$Q_{4.5}$	0	0	1	5	3	0	4.0
$Q_{4.6}$	0	0	1	5	2	1	3.4

7.1. Video Quality

To evaluate the quality of the videos, we propose four questions about four different aspects, namely: (i) how an expert perceives the similarity between the videos automatically generated using our method and those created by a teacher; (ii) the perception of the experts about the visual and textual explanations generated and (iii) the perception of the experts with respect to the usefulness of the videos to help users in the process of understanding the execution of a HNN. The questions are as follows: ($Q_{1.1}$) *"The automatically generated explainer video is close to or likely to the videos that could be created by a human (a teacher)"*; ($Q_{1.2}$) *"The automatically generated explainer video shows useful visual explanations to understand the execution of a Hopfield network"*; ($Q_{1.3}$) *"The automatically generated explainer video shows useful textual explanations to understand the execution of a Hopfield network."*; ($Q_{1.4}$) *"The automatically generated explainer video presents the information in a way that favors the understanding of the execution of a Hopfield network"*. The questions, answers and results are shown in Table 1 and Figure 11.

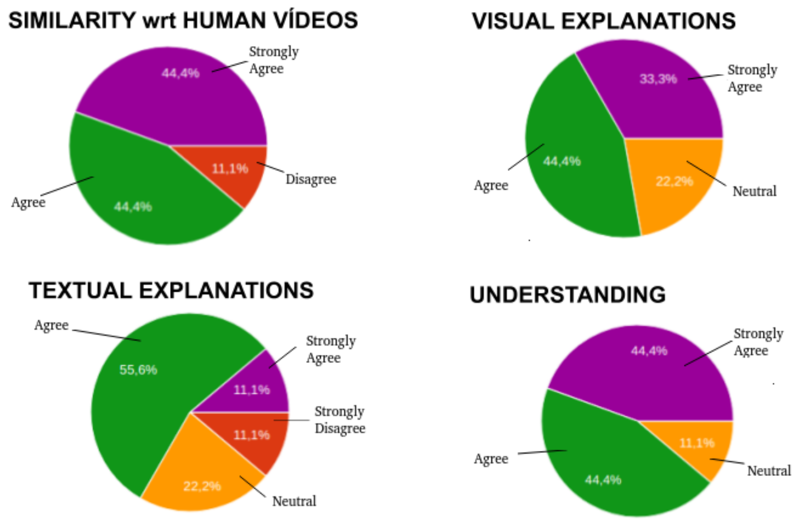


Figure 11. Percentages obtained from the answers of the experts for the questions about similarity regarding human, visual/textual explanations and understanding.

Now, we analyze the answer given by the experts. With respect to the first question: $Q_{1.1}$, 44.4% of the experts were strongly agree, 44.4% were agree and 11.1% were disagree.

With respect to the second one: $Q_{1.2}$, 33.3% of them were strongly agree, 44.4% were agree and 22.2% were disagree. With respect to the third question: $Q_{1.3}$, 11.1% were strongly agree, 55.6% were agree, 22.2% were neutral and 11.1% were disagree. Finally, with respect to the last question: ($Q_{1.4}$), 44.4% were strongly agree, 44.4% were agree and 11.1% were neutral. We can conclude that our explainer videos are a useful tool to get a better understanding of the execution of a HNN, but as the average of each aspect indicate (4.2, 4.1, 3.2 and 4.3) we will need to improve them in the corresponding aspects: visual and, above all, textual explanations.

7.2. Content Quality

Now, we evaluate the content of the frames; accuracy, balanced presentation of ideas and adequate level of detail are aspect to taken into account. We propose five questions about the mentioned aspects: ($Q_{2.1}$) *"The resource presents the information objectively and with good writing"*; ($Q_{2.2}$) *"The content does not present errors or omissions that could confuse or misinterpret its interpretation"*; ($Q_{2.3}$) *"The statements are supported by evidence or logical arguments"*; ($Q_{2.4}$) *"The information emphasizes the key points and the most significant ideas, with an appropriate level of detail"*; ($Q_{2.5}$) *"Is the language used in the explanation appropriate for you?"*.

We can conclude that the content of our explainer videos needs to be improved. This conclusion is directly related to the weaknesses detected in the previous step: textual explanations are poor. The worst result was identified in the question $Q_{2.5}$ in which some experts had doubts in choosing an answer. Therefore, all these aspects must be studied and improved in future works.

7.3. Motivation

The motivation is another important aspect to be evaluated. We can define it as the ability to generate interest in the public. We propose three questions to evaluate it: ($Q_{3.1}$) *"The resource offers a representation close to reality that stimulates the interest of the student"*; ($Q_{3.2}$) *"The duration of the content visualization favors the student's attention"*; ($Q_{3.3}$) *"Could students be motivated with this type of resource?"*.

We conclude that experts disagree with the duration of the videos and the current format of them. However, the most of the experts were agree with respect to the question about if the resource is motivating.

7.4. Design and Presentation

Finally, we evaluate two important aspects: design and presentation. In particular, we propose a questionnaire formed by six questions: ($Q_{4.1}$) *"The presentation of the video requires a minimum number of visual searches"*; ($Q_{4.2}$) *"The graphs and tables are clear, concise and without errors"*; ($Q_{4.3}$) *"Videos include narration"*; ($Q_{4.4}$) *"Paragraphs are headed by meaningful headings"*; ($Q_{4.5}$) *"The writing is clear, concise and without errors"*; ($Q_{4.6}$) *"The color and design are aesthetic"*.

We conclude that experts have a neutral opinion about the aspects related to the design/presentation of frames. It confirms that the narrative and storytelling is an important issue to solve in future works.

8. Conclusions and Future Work

A new framework based on data-to-text and software visualization to automatically generate explainer videos about the Hopfield neural networks' execution has been presented. A design based on experience and human-computer interface has been created and presented. A software architecture to build automatic video generation based on Computational Theory of Perceptions has been defined and explained in detail. Four modules form it:

1. Trace (for capturing the values of the data structures in each instant of time and other relevant information generated in the process of execution).

2. Frame generation/Local NLG (for creating slides containing texts and graphical representations about that).
3. Global NLG/Storytelling for the incorporation of texts which allows us to obtain a global communication.
4. Video generation (for generating a video in mp4 format from a set of images).

Each module has been implemented and tested using an object-based programming language called Java (version 11). We have focused on explaining the content determination and planning phases. Technical details about the implementation have been shown, the main algorithms and data structures implemented in this work have been explained. A real application for the generation of an explainer videos from HNN's execution has been designed, implemented and explained. Finally, we have positively evaluated that they are a useful tool to obtain a better understanding of the execution of a HNN, but we will need to improve the visual and, above all, textual explanations and the narrative in future works.

Therefore, as future work we would like to work on the following challenges:

1. To improve the video by incorporating sounds, voices and more advanced avatars.
2. To create an automatic style module for providing users with options about visual aspects of the video (colors, font, speed, so on).
3. To investigate new methods for generating more rich sentences and narratives.
4. To be able to work with large neural networks using intelligent analysis and automatic summarization of videos [62].
5. To employ links in which users can directly interact with the explainer video.

Special mention about future work is related to investigating how automatic explainer videos generation can be used as video tutorials for teaching-learning processes. The main reason is that videos online have become one of the most consumed digital resources, and millions of viewers watch them on YouTube and other platforms. For example, a tutorial on skin retouching in photoshop has been watched for more than a million times and a YouTube channel on photoshop tutorials has more than 170 thousand subscribers. In higher education, videos have also become an essential resource for students. They are integrated and employed as a fundamental part of traditional courses, and they serve as a cornerstone of many courses. On the other hand, they have become in the primary information-delivery mechanism in online classes. In fact, several experts have pointed out that e-learning is the future of education. In this regard, Global Industry Analysts projected e-learning will grow in the future [63]. Additionally, the current pandemic has driven the use of video tutorials which play an important role in the e-learning area. At the same time, it has been shown how the use of presentation software instructions on video has advantages with respect to those performed on paper [64]. Currently, video tutorials (recorded or live) are a fundamental resource for students because they provide them with another perspective which becomes the experience of learning in a more dynamic, practical and effective experience [65]. For all these reasons, automatic explainer video generation can be a useful approach for digital and classical learning in the future.

Author Contributions: The preprint version of this paper was originally written by C.R.-M. The revised version of this paper was written also by C.M.-A., A.S.-N. and C.V.-C. They have rewritten, revised and improved some part of this paper and they have designed the instrument of evaluation which is essential for the final version of this work. All authors have read and agreed to the published version of the manuscript.

Funding: University of the Bío-Bío. Vicerrectoría de Investigación. Facultad de Ciencias Empresariales. Departamento de Sistemas de Información.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

Acknowledgments: Facultad de Ciencias Empresariales and SOMOS Research Group.

Conflicts of Interest: The authors have no conflict of interest.

Appendix A. Algorithms (Pseudocode and Explanations)

Algorithm A1: Video Generation (Execution Trace T) returns a video in mp4.

```
V: Set of Frames                                          // A new empty set of frames is created
begin
    while not(end_file)                    // Rows of the log file are processed in the next bucle
     do
        Row[]=Read-Next-Line(Trace)     // Here, a new line is read (information needed for each frame)
        V.insert(Create-Slide(row[])) // A new frame is automatically generated and the result is inserted in the list of
          frames)
    end
    planning(V)                                                     // Frames are organized using a plan
    storytelling(V)                                                    // Storytelling is incorporated
    Vmp4=convert(V,mp4)                                         // Video is created and stored in MP4 file
end
return Vmp4                                      // The video is returned and it can be shown to the users
```

Algorithm A2: create-Slide(row-trace []).

```
begin
    graphics = createFrameBackGround-Style()                                    // Empty slide is created
    /* visual representations and texts are generated                                               */
    rfo=first-order-representations(info)          // First-order-representations are created and stored in rfo
    efo=explanations-NL-for(rfo)                          // Explanations are generated from the variable rfo
    rso=second-order-representations(rfo) // Second-order representations are created from rfo and stored in the variable
      rso
    eso=explanations-NL-for(rso)                          // Explanations are generated from the variable rso
    updatePositions(graphics,[rfo,efo,rso,eso])       // visual representations and texts are positioned in graphics
    paint(graphics,[rfo,efo,rso,eso])                            // visual representations and texts are painted
end
```

Algorithm A3: First-order-representation(info[]).

```
begin
    /* The first element of the list (info[0]) determines which type of frame will
       be generated using the function images-text (type, info)                      */
    switch info[0] do
        case "initialization" do
            images-text(initialization,info)          // initialization frames are created
            case "neurons" do
                images-text(neurons,info)                      // neurons frames are created
            end
            case "connections" do
                images-text(connections,info)              // connection frames are created
            end
            case "pattern" do
                images-text(pattern,info)                      // pattern frames are created
            end
            case "iterations" do
                images-text(iterations,info)               // iterations frames are created
            end
            case "conclusions" do
                images-text(conclusions,info)             // conclusions frames are created
            end
        end
        /* In this current version, several alternatives are possible (neurons,
           connections, pattern, iterations and conclusions).  Please note that
           depending on the type, the rest of the elements (info[i]) will be used to
           create an appropriate frame                                                  */
    end
end
```

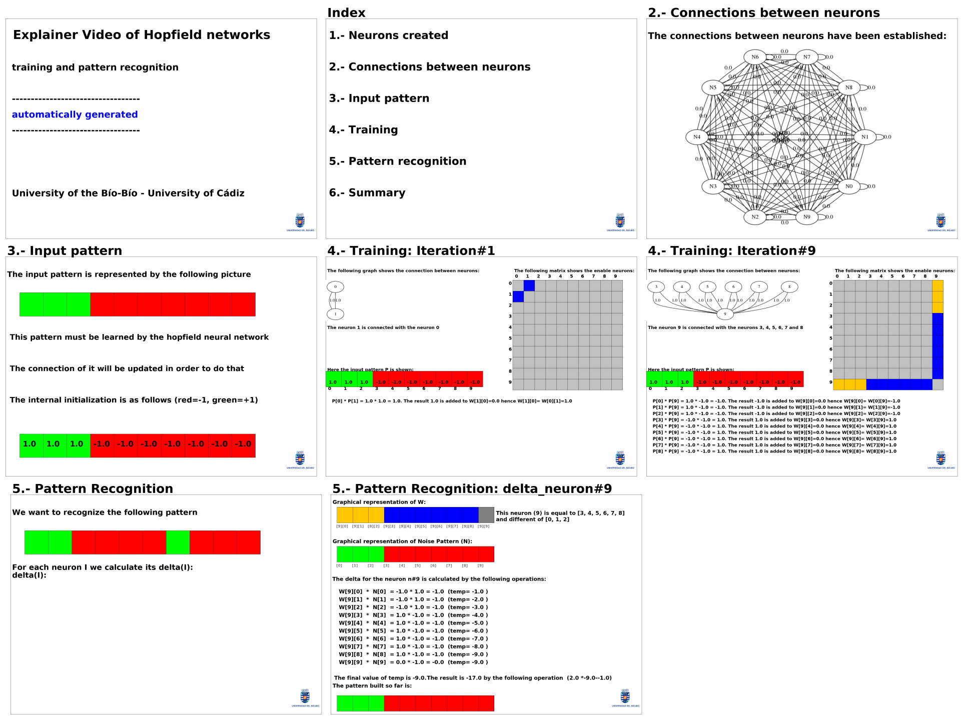
Algorithm A4: images-text-pattern(info[]).

```
begin
    /* The information is assigned to the corresponding variables (title, message1,
       message2, message3                                                          */
    title=info[1]
    message1=info[2]
    message2=info[3]
    message3=info[4]
    /* Variables are painted using 2D primitives.  The values (20, 60, so on) will
       depend on each type of frame                                                */
    setFont(fontTitle)
    drawString(title,20,60)
    setFont(fontBody);
    setColor(Color.black)
    drawString(message1,20,200)
    drawPattern(false,false,100,300,150)
    drawString(message2,20,600)
    drawString(message3,20,800)
    drawString(message4,20,1000)
    drawPattern(true,false,100,1200,150)
end
```

Algorithm A5: drawPattern(wn,wi,x,y,size)

```
begin
    /* A visual pattern has three important variables:  x, y and size, i.e., where
       is painted and how (size)                                                   */
    xcell=x
    ycell=y
    size=s
    for $i = 0$ to pattern.length do
        /* visual matrix with indexes is painted                                   */
        if wi=true then
            drawString(""+i,xcell+15,y+size+20);
        end
        /* red cell is painted if pattern[i] is -1.0.  Otherwise, green cell is
           painted                                                                 */
        if pattern[i]=="-1.0" then
            setColor(Color.red)
        end
        else
            setColor(Color.green)
        end
        /* 2D primitives to paint cells                                            */
        fillRect(xcell,ycell, size,size); setColor(Color.black); drawRect(xcell,ycell, size,size);
        /* visual matrix and values for each cell are painted                      */
        if wn=true then
            drawString(pattern[i],xcell+20,ycell+80);
        end
        xcell+=size;
    end
end
```

Appendix B. Frames Automatically Generated



Explainer Video of Hopfield networks

training and pattern recognition

----
automatically generated

University of the Bío-Bío - University of Cádiz

Index

1.- Neurons created
2.- Connections between neurons
3.- Input pattern
4.- Training
5.- Pattern recognition
6.- Summary

2.- Connections between neurons

The connections between neurons have been established:

3.- Input pattern

The input pattern is represented by the following picture

This pattern must be learned by the hopfield neural network

The connection of it will be updated in order to do that

The internal initialization is as follows (red=-1, green=+1)

1.0	1.0	1.0	-1.0	-1.0	-1.0	-1.0	-1.0	-1.0	-1.0

4.- Training: Iteration#1

The following graph shows the connection between neurons:

The neuron 1 is connected with the neuron 0

The following matrix shows the enable neurons:

Here the input pattern P is shown:

1.0	1.0	1.0	-1.0	-1.0	-1.0	-1.0	-1.0	-1.0	-1.0
0	1	2	3	4	5	6	7	8	9

```
P[0] * P[1] = 1.0 * 1.0 = 1.0. The result 1.0 is added to W[1][0]=0.0 hence W[1][0]= W[0][1]=1.0
```

4.- Training: Iteration#9

The following graph shows the connection between neurons:

The neuron 9 is connected with the neurons 3, 4, 5, 6, 7 and 8

The following matrix shows the enable neurons:

Here the input pattern P is shown:

1.0	1.0	1.0	-1.0	-1.0	-1.0	-1.0	-1.0	-1.0	-1.0
0	1	2	3	4	5	6	7	8	9

```
P[0] * P[9] = 1.0 * -1.0 = -1.0. The result -1.0 is added to W[9][0]=0.0 hence W[9][0]= W[0][9]=-1.0
P[1] * P[9] = 1.0 * -1.0 = -1.0. The result -1.0 is added to W[9][1]=0.0 hence W[9][1]= W[1][9]=-1.0
P[2] * P[9] = 1.0 * -1.0 = -1.0. The result -1.0 is added to W[9][2]=0.0 hence W[9][2]= W[2][9]=-1.0
P[3] * P[9] = -1.0 * -1.0 = 1.0. The result 1.0 is added to W[9][3]=0.0 hence W[9][3]= W[3][9]=1.0
P[4] * P[9] = -1.0 * -1.0 = 1.0. The result 1.0 is added to W[9][4]=0.0 hence W[9][4]= W[4][9]=1.0
P[5] * P[9] = -1.0 * -1.0 = 1.0. The result 1.0 is added to W[9][5]=0.0 hence W[9][5]= W[5][9]=1.0
P[6] * P[9] = -1.0 * -1.0 = 1.0. The result 1.0 is added to W[9][6]=0.0 hence W[9][6]= W[6][9]=1.0
P[7] * P[9] = -1.0 * -1.0 = 1.0. The result 1.0 is added to W[9][7]=0.0 hence W[9][7]= W[7][9]=1.0
P[8] * P[9] = -1.0 * -1.0 = 1.0. The result 1.0 is added to W[9][8]=0.0 hence W[9][8]= W[8][9]=1.0
```

5.- Pattern Recognition

We want to recognize the following pattern

For each neuron I we calculate its delta(I):
delta(I):

5.- Pattern Recognition: delta_neuron#9

Graphical representation of W:

This neuron (9) is equal to [3, 4, 5, 6, 7, 8]
and different of [0, 1, 2]

Graphical representation of Noise Pattern (N):

The delta for the neuron n#9 is calculated by the following operations:

```
W[9][0]  *  N[0]  = -1.0 * 1.0 = -1.0  (temp= -1.0 )
W[9][1]  *  N[1]  = -1.0 * 1.0 = -1.0  (temp= -2.0 )
W[9][2]  *  N[2]  = -1.0 * 1.0 = -1.0  (temp= -3.0 )
W[9][3]  *  N[3]  = 1.0 * -1.0 = -1.0  (temp= -4.0 )
W[9][4]  *  N[4]  = 1.0 * -1.0 = -1.0  (temp= -5.0 )
W[9][5]  *  N[5]  = 1.0 * -1.0 = -1.0  (temp= -6.0 )
W[9][6]  *  N[6]  = 1.0 * -1.0 = -1.0  (temp= -7.0 )
W[9][7]  *  N[7]  = 1.0 * -1.0 = -1.0  (temp= -8.0 )
W[9][8]  *  N[8]  = 1.0 * -1.0 = -1.0  (temp= -9.0 )
W[9][9]  *  N[9]  = 0.0 * -1.0 = -0.0  (temp= -9.0 )

 The final value of temp is -9.0.The result is -17.0 by the following operation  (2.0 *-9.0--1.0)
The pattern built so far is:
```

Figure A1. A subset of frames automaically generated.

References

1. Fluck, M.; McCarthy, D.R. Information is power? Transparency and fetishism in international relations. *Globalizations* **2019**, *16*, 1–16. [CrossRef]
2. Florini, A. *The Right to Know: Transparency for an Open World*; Columbia University Press: New York, NY, USA, 2007.
3. Castelvecchi, D. Can we open the black box of AI? *Nat. News* **2016**, *538*, 20. [CrossRef] [PubMed]
4. Datta, A.; Sen, S.; Zick, Y. Algorithmic transparency via quantitative input influence: Theory and experiments with learning systems. In Proceedings of the 2016 IEEE Symposium on Security and Privacy (SP), San Jose, CA, USA, 22–26 May 2016; pp. 598–617.
5. Diakopoulos, N.; Koliska, M. Algorithmic transparency in the news media. *Digit. J.* **2017**, *5*, 809–828. [CrossRef]
6. Kroll, J.A. Accountable Algorithms. Ph.D. Dissertation, Princeton University, Princeton, NJ, USA, 2015.
7. Kemper, J.; Kolkman, D. Transparent to whom? No algorithmic accountability without a critical audience. *Inf. Commun. Soc.* **2019**, *22*, 2081–2096. [CrossRef]
8. Veale, M. Logics and practices of transparency and opacity in real-world applications of public sector machine learning. *arXiv* **2017**, arXiv:1706.09249.
9. Belle, V. Symbolic logic meets machine learning: A brief survey in infinite domains. In *International Conference on Scalable Uncertainty Management*; Springer: Cham, Switzerland, 2020; pp. 3–16.
10. Bücker, M.; Szepannek, G.; Gosiewska, A.; Biecek, P. Transparency, Auditability and eXplainability of Machine Learning Models in Credit Scoring. *arXiv* **2020**, arXiv:2009.13384.
11. Confalonieri, R.; Coba, L.; Wagner, B.; Besold, T.R. A historical perspective of explainable Artificial Intelligence. *Wiley Interdiscip. Rev. Data Min. Knowl. Discov.* **2021**, *11*, e1391. [CrossRef]
12. Confalonieri, R.; Weyde, T.; Besold, T.R.; del Prado Martín, F.M. Using ontologies to enhance human understandability of global post-hoc explanations of black-box models. *Artif. Intell.* **2021**, *296*, 103471. [CrossRef]
13. Gunning, D. Explainable artificial intelligence (xai). *Def. Adv. Res. Proj. Agency Nd Web* **2017**, *2*, 1–18. [CrossRef]
14. Jin, W.; Fan, J.; Gromala, D.; Pasquier, P.; Hamarneh, G. EUCA: A Practical Prototyping Framework towards End-User-Centered Explainable Artificial Intelligence. *arXiv* **2021**, arXiv:2102.02437.
15. Rudin, C.; Chen, C.; Chen, Z.; Huang, H.; Semenova, L.; Zhong, C. Interpretable machine learning: Fundamental principles and 10 grand challenges. *arXiv* **2021**, arXiv:2103.11251.
16. Stepin, I.; Alonso, J.M.; Catala, A.; Pereira-Fariña, M. A survey of contrastive and counterfactual explanation generation methods for explainable artificial intelligence. *IEEE Access* **2021**, *9*, 11974–12001. [CrossRef]

17. Darlington, K. Explainable AI Systems: Understanding the Decisions of the Machines. In *BBVA Open Mind*; BBVA; Spain; 2017. Available online: http://www.bbvaopenmind.com/en/technology/artificial-intelligence/explainable-ai-systems-understanding-the-decisions-of-the-machines/ (accessed on 19 June 2021).
18. Burkart, N.; Huber, M.F. A Survey on the Explainability of Supervised Machine Learning. *J. Artif. Intell. Res. X* **2020**, *70*, 245–317. [CrossRef]
19. Goodman, B.; Flaxman, S. European Union regulations on algorithmic decision-making and a "right to explanation". *AI Mag.* **2017**, *38*, 50–57. [CrossRef]
20. Antwarg, L.; Shapira, B.; Rokach, L. Explaining anomalies detected by autoencoders using SHAP. *arXiv* **2019**, arXiv:1903.02407.
21. Ribeiro, M.T.; Singh, S.; Guestrin, C. "Why should i trust you?" Explaining the predictions of any classifier. In Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, San Francisco, CA, USA, 13–17 August 2016; pp. 1135–1144.
22. Shrikumar, A.; Greenside, P.; Kundaje, A. Learning important features through propagating activation differences. In Proceedings of the International Conference on Machine Learning, Sydney, Australia, 6–11 August 2017; pp. 3145–3153.
23. Duch, W. Coloring black boxes: Visualization of neural network decisions. In Proceedings of the International Joint Conference on Neural Networks, Portland, OR, USA, 20–24 July 2003; Volume 3, pp. 1735–1740.
24. Morch, N.J.; Kjems, U.; Hansen, L.K.; Svarer, C.; Law, I.; Lautrup, B.; Rehm, K. Visualization of neural networks using saliency maps. In Proceedings of the ICNN'95-International Conference on Neural Networks, Perth, WA, Australia, 27 November–1 December 1995; Volume 4, pp. 2085–2090.
25. Samek, W.; Binder, A.; Montavon, G.; Lapuschkin, S.; Müller, K.R. Evaluating the visualization of what a deep neural network has learned. *IEEE Trans. Neural Netw. Learn. Syst.* **2016**, *28*, 2660–2673. [CrossRef] [PubMed]
26. Tzeng, F.Y.; Ma, K.L. *Opening the Black Box-Data Driven Visualization of Neural Networks*; IEEE: New York, NY, USA, 2005; pp. 383–390.
27. Chen, J.; Hu, K.; Yu, Y.; Chen, Z.; Xuan, Q.; Liu, Y.; Filkov, V. Software visualization and deep transfer learning for effective software defect prediction. In Proceedings of the ACM/IEEE 42nd International Conference on Software Engineering, Seoul, Korea, 24 June–16 July 2020; pp. 578–589.
28. Li, M.; Zhao, Z.; Scheidegger, C. Visualizing Neural Networks with the Grand Tour. *Distill* **2020**, *5*, e25. [CrossRef]
29. Archambault, D.; Purchase, H.; Pinaud, B. Animation, small multiples, and the effect of mental map preservation in dynamic graphs. *IEEE Trans. Vis. Comput. Graph.* **2010**, *17*, 539–552. [CrossRef]
30. Tversky, B.; Morrison, J.B.; Betrancourt, M. Animation: Can it facilitate? *Int. J. Hum. Comput. Stud.* **2002**, *57*, 247–262. [CrossRef]
31. Alonso, J.M.; Barro, S. Interactive Natural Language Technology for Explainable Artificial Intelligence. In Proceedings of the Workshop on Foundations of Trustworthy AI integrating Learning, Optimisation and Reasoning, at the European Conference on Artificial Intelligence, Virtual Event, 4–5 September 2020.
32. Bloch, G.R. From concepts to film sequences. In *User-Oriented Content-Based Text and Image Handling*; LE CENTRE DE HAUTES ETUDES INTERNATIONALES D'INFORMATIQUE DOCUMENTAIRE 36bis rue Ballu: Paris, France 1988; pp. 760–767.
33. Rieber, L.P. *Computers Graphics and Learning*; Brown & Benchmark Pub: Hull, GA, USA, 1994.
34. Mayer, R.E. Systematic thinking fostered by illustrations in scientific text. *J. Educ. Psychol.* **1989**, *81*, 240. [CrossRef]
35. Mayer, R.E.; Gallini, J.K. When is an illustration worth ten thousand words? *J. Educ. Psychol.* **1990**, *82*, 715. [CrossRef]
36. Davis, M. Knowledge representation for video. In Proceedings of the AAAI-94 Proceedings, Thetwelfth National Conference on Artificial Intelligence, Seattle, WA, USA, 1994; pp. 120–127.
37. Glowacz, A. Fault diagnosis of electric impact drills using thermal imaging. *Measurement* **2021**, *171*, 108815. [CrossRef]
38. Hong, K.; Uh, Y.; Byun, H. ArrowGAN: Learning to Generate Videos by Learning Arrow of Time. *arXiv* **2021**, arXiv:2101.03710.
39. Truong, A.; Chi, P.; Salesin, D.; Essa, I.; Agrawala, M. Automatic Generation of Two-Level Hierarchical Tutorials from Instructional Makeup Videos. In Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems, Yokohama, Japan, 7–17 May 2021; pp. 1–16.
40. Richardson, K.; Zarrieß, S.; Kuhn, J. The code2text challenge: Text generation in source code libraries. *arXiv* **2017**, arXiv:1708.00098.
41. Cheng, K.S.; Lin, J.S.; Mao, C.W. The application of competitive Hopfield neural network to medical image segmentation. *IEEE Trans. Med. Imaging* **1996**, *15*, 560–567. [CrossRef]
42. Hopfield, J.J. Neural networks and physical systems with emergent collective computational abilities. *Proc. Natl. Acad. Sci. USA* **1982**, *79*, 2554–2558. [CrossRef]
43. Ramsauer, H.; Schäfl, B.; Lehner, J.; Seidl, P.; Widrich, M.; Adler, T.; Gruber, L.; Holzleitner, M.; Pavlović, M.; Sandve, G.K.; et al. Hopfield networks is all you need. *arXiv* **2020**, arXiv:2008.02217.
44. Ramos-Soto, A.; Bugarín, A.J.; Barro, S.; Taboada, J. Linguistic Descriptions for Automatic Generation of Textual Short-Term Weather Forecasts on Real Prediction Data. *IEEE Trans. Fuzzy Syst.* **2015**, *23*, 44–57. [CrossRef]
45. Reiter, E.; Sripada, S.; Hunter, J.; Yu, J.; Davy, I. Choosing words in computer-generated weather forecasts. *Artif. Intell.* **2005**, *167*, 137–169. [CrossRef]
46. Sripada, S.; Reiter, E.; Davy, I. SumTime-Mousam: Configurable marine weather forecast generator. *Expert Update* **2003**, *6*, 4–10.
47. Sripada, S.; Burnett, N.; Turner, R.; Mastin, J.; Evans, D. A Case Study: NLG meeting Weather Industry Demand for Quality and Quantity of Textual Weather Forecasts. In Proceedings of the 8th International Natural Language Generation Conference (INLG), Philadelphia, PA, USA, 19–21 June 2014; pp. 1–5.

48. Zadeh, L.A. From computing with numbers to computing with words. From manipulation of measurements to manipulation of perceptions. *IEEE Trans. Circuits Syst. I Fundam. Theory Appl.* **1999**, *46*, 105–119. [CrossRef]
49. Zadeh, L.A. Toward human level machine intelligence-is it achievable? the need for a paradigm shift. *IEEE Comput. Intell. Mag.* **2008**, *3*, 11–22. [CrossRef]
50. Trivino, G.; Sugeno, M. Towards linguistic descriptions of phenomena. *Int. J. Approx. Reason.* **2013**, *54*, 22–34. [CrossRef]
51. Eciolaza, L.; Pereira-Fari na, M.; Trivino, G. Automatic linguistic reporting in driving simulation environments. *Appl. Soft Comput.* **2013**, *13*, 3956–3967. [CrossRef]
52. Rubio-Manzano, C.; Trivino, G. Improving player experience in computer games by using players' behavior analysis and linguistic descriptions. *Int. J. Hum. Comput. Stud.* **2016**, *95*, 27–38. [CrossRef]
53. Conde-Clemente, P.; Alonso, J.M.; Trivino, G. Toward automatic generation of linguistic advice for saving energy at home. *Soft Comput.* **2016**, *22*, 345–359. [CrossRef]
54. Forsdale, J.R.; Forsdale, L. Film literacy. *J. Univ. Film Prod. Assoc.* **1996**, *18*, 9–27.
55. Ellson, J.; Gansner, E.; Koutsofios, L.; North, S.C.; Woodhull, G. Graphviz—Open source graph drawing tools. In *International Symposium on Graph Drawing*; Springer: Berlin/Heidelberg, Germany, 2001; pp. 483–484.
56. Trivino, G.; Sobrino, A. Human Perceptions versus Computational Perceptions in Computational Theory of Perceptions. In Proceedings of the IFSA/EUSFLAT Conference, Lisbon, Portugal, 20–24 July 2009; pp. 327–332.
57. Halliday, M.A.K.; Matthiessen, C. *Construing Experience through Meaning: A Language-Based Approach to Cognition*; Bloomsbury Publishing: London, UK, 2006.
58. Nevatia, R.; Hobbs, J.; Bolles, B. An ontology for video event representation. In Proceedings of the 2004 Conference on Computer Vision and Pattern Recognition Workshop, Washington, DC, USA, 27 June–2 July 2004; p. 119.
59. Anderson, C.; Duarte, N. How to give a killer presentation. *Harv. Bus. Rev.* **2013**, *91*, 121–125.
60. Riedl, M.O.; Young, R.M. Narrative planning: Balancing plot and character. *J. Artif. Intell. Res.* **2010**, *39*, 217–268. [CrossRef]
61. Nesbit, J.C.; Belfer, K.; Leacock, T. Learning Object Review Instrument (LORI). 2003. Available online: http://edutechwiki.unige.ch/en/Learning_Object_Review_Instrument (accessed on 21 June 2021).
62. He, L.; Sanocki, E.; Gupta, A.; Grudin, J. Auto-summarization of audio-video presentations. In Proceedings of the Seventh ACM International Conference on MULTIMEDIA (Part 1), Orlando, FL, USA, 30 October 1999; pp. 489–498.
63. McCue, T. *E Learning Climbing to 325 Billion Dollars by 2025 uf Canvas Absorb Schoology Moodle*. 2018. Available online: https://www.forbes.com/sites/tjmccue/2018/07/31/e-learning-climbing-to-325-billion-by-2025-uf-canvas-absorb-schoology-moodle/ (accessed on 21 June 2021).
64. van der Meij, H.; Van Der Meij, J. A comparison of paper-based and video tutorials for software learning. *Comput. Educ.* **2014**, *48*, 150–159. [CrossRef]
65. Windermere, A. *What Is the Importance of Video Tutorials to Students*. 2019. Available online: https://work.chron.com/importance-video-tutorials-students-16633.html (accessed on 21 June 2021).