

Predición de preferencia de usuario mediante técnicas de *soft computing*

por Iria María Santos López

Tese doutoral – Universidade da Coruña – 2021

Titor: Juan Jesús Romero Cardalda

Directores: Adrián Carballal Mato e Juan Jesús Romero Cardalda

Programa de Doutoramento en Tecnoloxías da Información e as Comunicaci3ns



UNIVERSIDADE DA CORUÑA

Autora: Iria María Santos López

Titor: Juan Jesús Romero Cardalda

Directores: Adrián Carballal Mato e Juan Jesús Romero Cardalda

Universidade da Coruña

D. **Adrián Carballal Mato** e D. **Juan Jesús Romero Cardalda**, profesores da Universidade da Coruña do Departamento de Ciencias da Computación e Tecnoloxías da Información como directores da presente Tese Doutoral:

FAN CONSTAR QUE:

A tese titulada: “**Predición de preferencia de usuario mediante técnicas de soft computing**” realizada por Dna. Iria María Santos López, baixo a miña dirección no Departamento de Ciencias da Computación e Tecnoloxías da Información da Universidade da Coruña cumpre cos requisitos para optar ao grao de Doutora.

E para que así conste asinan a presente a 25 de maio de 2021.

Adrián Carballal Mato | Juan Jesús Romero Cardalda | Iria María Santos López

Autora: **Iria María Santos López**
Titor: **Juan Jesús Romero Cardalda**
Directores: **Adrián Carballal Mato** e **Juan Jesús Romero Cardalda**
Universidade da Coruña

Ao meu pai, á miña nai e ao resto de familia,

tanto á de sangue como á escollida.

Polos que están e os que estiveron.

En memoria de Milagros.

Autora: Iria María Santos López
Titor: Juan Jesús Romero Cardalda
Directores: Adrián Carballal Mato e Juan Jesús Romero Cardalda
Universidade da Coruña

Agradecementos

É de ben nacido ser agradecido.

O Instituto de Saúde Carlos III do Plan Nacional Español de Investigación e Innovación Científica e Técnica 2013-2016 e os Fondos Europeos de Desenvolvemento Rexional (FEDER) “Un xeito de construír Europa” apoian este traballo a través do “Proxecto colaborativo en integración de datos xenómicos (CICLOGEN)” PI17/01826. Este traballo conta tamén co apoio da Dirección Xeral de Cultura, Educación e Ordenación Universitaria da Xunta de Galicia (Refs. ED431D, ED431G/01 2017/16), a “Rede galega para a investigación do cancro colorrectal” (Ref. ED431D 2017/23) e Grupos de Referencia Competitivos (Ref. ED431C 2018/49). Este traballo tamén foi apoiado polo CITIC, como Centro de Investigación do Sistema Universitario de Galicia, financiado pola Consellería de Educación, Universidade e Formación Profesional da Xunta de Galicia a través do Fondo Europeo de Desenvolvemento Rexional (FEDER) cun 80%, Programa Operativo FEDER Galicia 2014-2020, e o 20% restante pola Secretaría Xeral de Universidades (Ref. ED431G 2019/01). Por outra banda, a instalación única BIOCAI (UNLC08-1E-002 UNLC13-13-3503) foi financiada polo Ministerio de Economía e Competitividade español e os Fondos Europeos de Desenvolvemento Rexional (FEDER). Agradecementos tamén ao apoio de NVIDIA Corporation pola doazón da GPU Tesla K40 empregada nesta investigación.

Son moitos os agradecementos que debo incluír nesta tese doutoral, xa que sen a axuda e apoio dun gran número de persoas non tería chegado ata aquí. O primeiro deles, sen dúbida, o meu director e titor de tese, **Juan Jesús Romero Cardalda**, que foi quen me guiou e me deu as nocións necesarias para poder levar o traballo adiante. A súa paciencia e axuda foi un punto fundamental á hora de poder finalizar este proxecto. El foi quen de mostrarme o máis bonito da **IA** e a súa aplicación nas artes visuais. Ademais, a súa paixón e ganas no seu labor de investigación servíronme como referente e así seguirá sendo no futuro. Grazas de todo corazón.

Gustaríame tamén mostrar o meu agradecemento a **todos os coautores dos artigos**, que forman parte desta tese por compendio, posto que sen a súa participación e colaboración, non sería posible finalizala a día de hoxe. Entre os distintos coautores, quero facer unha especial mención a **Adrián Carballal Mato**, tamén director desta tese, por colaborar de forma activa no seguimento e por tratarse

Autora: **Iria María Santos López**

Titor: **Juan Jesús Romero Cardalda**

Directores: **Adrián Carballal Mato** e **Juan Jesús Romero Cardalda**

Universidade da Coruña

dunha peza fundamental no desenvolvemento dos tres artigos de investigación resultantes dos estudos.

Tamén desexo agradecerlle a **todos os compañeiros do Grupo de Investigación RNASA-IMEDIR**, que en maior ou menor medida, me brindaron apoio ao longo do meu proxecto e, sobre todo, aos integrantes da liña de investigación de **computación evolutiva e estética computacional** cos cales teño un tratamento máis próximo e que son fonte de coñecemento e axuda en todo momento. Tamén quero facer unha mención especial a **Alejandro Celestino Pazos Sierra**, xa que como catedrático da Universidade da Coruña, director do Departamento de Ciencias da Computación e Tecnoloxías da Información (no cal se engloba a miña tese de doutoramento), e fundador e director do grupo de investigación mencionado previamente, fai un labor excelente e sempre está pendente de que os doutorandos poidamos tomar parte activa en diversas investigacións.

Xunto con isto, agradecer aos participantes da *spin off* da Universidade da Coruña, **Artificial Intelligence INDESTIA S.L.**, xa presentes nas mencións anteriores, pero onde se inclúe tamén **Manuel V. Pérez García**, que dende que entrou a formar parte como CEO de **PhotoLike** non deixou en ningún momento de achegar valor e axuda en todo o que ten na súa man.

Mostrar tamén o meu agradecemento a **Manuel Álvarez Díaz** e a **Julian Alfonso Dorado de la Calle**, que dende a **CAPD** estiveron sempre presentes e deron resposta rápida a todas as consultas, que se lles foron dirixidas, axudando ao correcto desenvolvemento do proxecto; aos **integrantes da Unidade de Apoio aos Departamentos de Investigación (UADI) da Facultade de Informática**, por estaren dispoñibles en todo momento para resolver problemas e responder cuestións; e á **Dirección da Escola Internacional de Doutoramento**.

A **Forsythe et al.** (2011) pola cesión do seu conxunto de datos para poder avanzar na investigación con base nos seus resultados previos, xa que sen a súa colaboración non podería terse levado a cabo o experimento final correctamente.

Aos membros do **Servizo de Normalización Lingüística da Universidade da Coruña**, xa que grazas ás súas correccións ortográficas e/ou gramaticais esta tese ten mellor nivel, que a versión orixinal sen corrixir.

Agradézolle tamén aos **compañeiros, profesores e orientadores**, que tiveron durante todos os anos previos a esta tese, dende os meus **inicios**, pasando pola

Autora: **Iria María Santos López**

Titor: **Juan Jesús Romero Cardalda**

Directores: **Adrián Carballal Mato** e **Juan Jesús Romero Cardalda**

Universidade da Coruña

ESO, bacharelato, ciclo superior, o grao e os másters, porque se encargaron de darme o empurrón que necesitaba en cada momento para poder avanzar. Primeiro aos que tiñan confianza en min, pero tamén aos que se encargaron de quitarme valor ou pensaron que non chegaría lonxe, porque realmente tamén foron parte do impulso que necesitaba para conseguir as miñas metas.

Por último, a meus pais, **Manuel J. Santos Mareque e M^a Eva López Campelo**, porque son os meus piares, déronme a vida e danme as forzas día tras día. Ao resto de **familia e amigos**, que están pendentes dos meus progresos e se senten orgullosos dos meus logros. Así como ao meu mozo, **José R. Díaz Gómez**, que estivo ao meu carón durante todo o proceso desta tese e tivo que aturar de primeira man os altibaixos; éche un anaco de paciencia metido no corpo dun ser humano, pero non sei se se decatou aínda. E sendo un pouco ególatra, tamén grazas **a min**, por converter o que ao principio só eran soños nunha enorme realidade.

Autora: **Iria María Santos López**

Titor: **Juan Jesús Romero Cardalda**

Directores: **Adrián Carballal Mato e Juan Jesús Romero Cardalda**

Universidade da Coruña

Prefacio

Esta tese preséntase na modalidade de “Compendio de artigos de investigación”, cumprindo os requisitos esixidos no artigo **41 do Regulamento de estudos de doutoramento da Universidade da Coruña**.

No presente documento, séguense as indicacións xerais recollidas no devandito regulamento, de modo que se achega unha **introdución**, que inclúe unha **xustificación** de maneira razoada sobre a unidade e coherencia temática e metodolóxica da tese en relación cos **obxectivos** que se busca alcanzar, así como unha **discusión xeral**, as **conclusións** e a **bibliografía común**.

A continuación, as referencias dos artigos publicados en revistas indexadas no JOURNAL CITATIONS REPORT (JCR), base de datos de recoñecido prestixio internacional.

Artigo 1: Santos, I., Castro, L., Rodríguez-Fernandez, N., Torrente-Patiño, Á., Carballal, A. *Artificial Neural Networks and Deep Learning in the Visual Arts: a review*. *Neural Comput & Applic* (2021). <https://doi.org/10.1007/s00521-020-05565-4>

[Q1; JCR ranking 2019: COMPUTER SCIENCE, ARTIFICIAL INTELLIGENCE 23/137; Impact Factor 4.774].

Artigo 2: Carballal, A., Cedron, F., Santos, I., Santos A., Romero J., *Minimal neural network topology optimization for aesthetic classification*. *Neural Comput & Applic* (2021). <https://doi.org/10.1007/s00521-020-05550-x>

[Q1; JCR ranking 2019: COMPUTER SCIENCE, ARTIFICIAL INTELLIGENCE 23/137; Impact Factor 4.774].

Artigo 3: Carballal, A., Fernandez-Lozano, C., Rodriguez-Fernandez, N., Santos, I., & Romero, J. (2020). *Comparison of Outlier-Tolerant Models for Measuring Visual Complexity*. *Entropy*, 22(4), 488. <https://doi.org/10.3390/e22040488>

[Q2; JCR ranking 2019: MATHEMATICS, INTERDISCIPLINARY APPLICATIONS 28/106; MULTIDISCIPLINARY SCIENCES 31/71; Impact Factor 2.462].

Autora: Iria María Santos López

Titor: Juan Jesús Romero Cardalda

Directores: Adrián Carballal Mato e Juan Jesús Romero Cardalda

Universidade da Coruña

Índice

Agradecementos.....	7
Prefacio.....	10
Resumo.....	14
Abstract.....	15
Resumen.....	16
1 Introducción.....	18
1.1 Belas artes e tarefas artísticas.....	20
1.2 IA, <i>Machine Learning</i> e <i>Deep Learning</i>	25
1.3 Técnicas adaptativas.....	29
1.4 IA aplicada a tarefas artísticas.....	31
2 Xustificación e obxectivos.....	33
2.1 Metodoloxía.....	34
2.2 Artigos.....	35
3 <i>Artificial Neural Networks and Deep Learning in the Visual Arts: a review</i>	37
3.1 Detección de obxectos.....	38
3.2 Clasificación segundo o autor/a ou estilo.....	40
3.3 Clasificación baseada na calidade, complexidade e características visuais.....	41
3.4 Avaliación da calidade fotográfica e o valor estético.....	43
3.5 Transferencia de estilo e xeración pictórica ou reconstrución.....	44
4 Minimal Neural Network Topology Optimization for Aesthetic Classification.....	48
4.1 <i>Dataset</i> e traballos previos.....	49
4.2 Métricas empregadas.....	52
4.3 Arquitectura proposta.....	54
5 Comparison of Outlier-Tolerant Models for Measuring Visual Complexity.....	60
5.1 <i>Dataset</i> e traballos previos.....	61
5.2 Modelos estudados.....	65
6 Conclusións.....	69
Referencias.....	71
ANEXOS (Publicacións).....	83

Autora: Iria María Santos López

Titor: Juan Jesús Romero Cardalda

Directores: Adrián Carballal Mato e Juan Jesús Romero Cardalda

Universidade da Coruña

Resumo

O contido desta tese por compendio é a agrupación de tres artigos de investigación publicados en revistas de prestixio, que abordan a necesidade e a forma de mellorar os métodos de predicción de preferencia estética de usuario mediante técnicas de *soft computing*.

Realízase un amplo estado da arte do uso de redes de neuronas artificiais e *deep learning*. Este estudo mostra que existen sistemas baseados en redes de neuronas capaces de realizaren tarefas artísticas con diferente grao de obxectividade. Dende a detección dun obxecto nunha obra pictórica ata a creación de imaxes, pasando pola clasificación segundo o estilo artístico ou autor/a, ou a estimación de calidade e valor estético. Tamén mostra que nos últimos anos se están a realizar máis traballos en tarefas máis complexas como a creación de imaxes, en gran medida grazas ao uso de técnicas de *deep learning* caso das *Convolutional Neural Network* (CNN) e as *Generative Adversarial Network* (GAN).

A partir desta base, fórmulase o uso de sistemas baseados en redes neuronais para dúas tarefas relevantes no ámbito da predicción de preferencia estética. Dunha banda, emprégase un sistema de redes de neuronas artificiais para predicción estética, utilizando un *dataset* explorado polo estado da arte. Non se busca só un erro baixo na predicción senón tamén unha rede cuxa topoloxía sexa mínima. Analízanse os resultados extraendo conclusións sobre a información mínima relevante para realizar esta tarefa altamente subxectiva e complexa.

Pola outra, analízanse diferentes alternativas para outra tarefa altamente relacionada coa percepción estética: a percepción de complexidade visual. Existen numerosos estudos psicolóxicos que propoñen unha relación directa entre complexidade e valor estético. Proponse buscar un método de *machine learning*, que obteña mellor predicción deste valor. Tamén se realiza unha análise dos *outlayers*, co fin de comprender mellor os procesos realizados polo mecanismo de predicción.

Autora: **Iria María Santos López**

Titor: **Juan Jesús Romero Cardalda**

Directores: **Adrián Carballal Mato** e **Juan Jesús Romero Cardalda**

Universidade da Coruña

Abstract

The content of this Thesis by Compendium is the grouping of three research articles published in prestigious journals, which shows the need and how to improve the methods of predicting user aesthetic preference using soft computing techniques.

An extensive state of the art of the use of artificial neural networks and deep learning is performed. This study shows that there are systems based on neural networks capable of performing artistic tasks with varying degrees of objectivity. From the detection of an object in a pictorial work to the creation of images, the classification according to the artistic style or author, or the estimation of quality and aesthetic value. It also shows that in recent years more work is being done on more complex tasks such as image creation, largely thanks to the use of deep learning techniques such as Convolutional Neural Network (*CNN*) and Generative Adversarial Network (*GAN*).

From this basis, the use of neural network-based systems is formulated for two relevant tasks in the field of aesthetic preference prediction. On the one hand, a system of artificial neural networks is used for aesthetic prediction, using a dataset explored by the state of the art. Not only a low error in the prediction is sought but also a network whose topology is minimal. The results are analyzed by drawing conclusions about the minimum information relevant to perform this highly subjective and complex task.

On the other hand, different alternatives for another task highly related to aesthetic perception are analyzed: the perception of visual complexity. There are numerous psychological studies that propose a direct relationship between complexity and aesthetic value. It is proposed to look for a machine learning method, which obtains better prediction of this value. An analysis of the outliers is also performed, in order to better understand the processes performed by the prediction mechanism.

Autora: Iria María Santos López

Titor: Juan Jesús Romero Cardalda

Directores: Adrián Carballal Mato e Juan Jesús Romero Cardalda

Universidade da Coruña

Resumen

El contenido de esta Tesis por Compendio es la agrupación de tres artículos de investigación publicados en revistas de prestigio, que abordan la necesidad y la forma de mejorar los métodos de predicción de preferencia estética de usuario mediante técnicas de *soft computing*.

Se realiza un amplio estado del arte del uso de redes de neuronal artificiales y *deep learning*. Este estudio muestra que existen sistemas basados en redes de neuronas capaces de realizar tareas artísticas con diferente grado de objetividad. Desde la detección de un objeto en una obra pictórica hasta la creación de imágenes, pasando por la clasificación según el estilo artístico o autor, o la estimación de calidad y valor estético. También se muestra que en los últimos años se están realizando más trabajos en tareas más complejas como la creación de imágenes, en gran medida gracias al uso de técnicas de *deep learning* como las *Convolutional Neural Networks (CNN)* y las *Generative Adversarial Networks (GAN)*.

A partir de esta base, se plantea el uso de sistemas basados en redes neuronales para dos tareas relevantes en el ámbito de la predicción de preferencia estética. Por un lado, se emplea un sistema de redes de neuronas artificiales para predicción estética, utilizando un *dataset* explorado por el estado del arte. No sólo se busca un error bajo en la predicción, sino también una red cuya topología sea mínima. Se analizan los resultados extrayendo conclusiones sobre la información mínima relevante para realizar esta tarea altamente subjetiva y compleja.

Por otro lado, se analizan diferentes alternativas para otra tarea altamente relacionada con percepción estética: la percepción de complejidad visual. Existen numerosos estudios psicológicos que proponen una relación directa entre complejidad y valor estético. Se propone buscar un método de *machine learning*, que obtenga mejor predicción de este valor. También se realiza un análisis de los *outlayers*, con el fin de comprender mejor los procesos realizados por el mecanismo de predicción.

Autora: **Iria María Santos López**

Titor: **Juan Jesús Romero Cardalda**

Directores: **Adrián Carballal Mato** e **Juan Jesús Romero Cardalda**

Universidade da Coruña

1 Introducción

Existen poucas tarefas que sexan tan características do ser humano como o son as artísticas. A **intelixencia artificial (IA)** utilízase a día de hoxe en tarefas artísticas obxectivas, pero a investigación relacionada co seu uso en aplicacións de maior subxectividade aínda é escaso. O valor dos traballos artísticos dependen de diversos factores como a cultura, os gustos do observador ou o tempo. Resulta especialmente atractivo modelizar estes aspectos creativos mediante **IA**, pero o avance durante moitos anos neste campo foi lento. Nos anos noventa, en cambio, a área tivo un auxe importante co emprego de técnicas de *soft computing* (computación evolutiva e redes neuronais, sobre todo). Durante os últimos anos, houbo un número crecente de artigos propostos sobre o uso de técnicas de **IA** para a creación e a análise de obras de arte en distintos campos: música, poesía, danza, escultura, arquitectura, fotografía, pintura... Incluso existen a día de hoxe conferencias anuais dedicadas a estes ámbitos como a **International Conference on Artificial Intelligence on Music, Art and Design (EvoMUSART)** e a **International Conference on Computational Creativity (ICCC)**, entre outras, e publícanse números especiais en diferentes xornais (**ALIFE** -Q1-, **Leonardo** -revista de referencia en arte y tecnología-, **Mathematics and the Arts**, **Entropy** -Q2-, **Neural Computing & Applications** -Q1- ou **Complexity** -Q1-) ao redor desta temática. Destaca tamén o libro editado por **Romero e Machado** (2008) cuxo título é “The Art of Artificial Evolution: A Handbook on Evolutionary Art and Music”.

Esta tese de doutoramento componse de tres artigos científicos publicados en dous dos xornais de referencia mencionados con anterioridade (**Neural Computing & Applications**, e **Entropy**).

Estes tres artigos abordan o uso de técnicas de *soft computing* para a mellora dos métodos de predición de preferencia estética de usuario, mediante tres achegas complementarias. Os artigos son os seguintes:

- 1 Santos, I., Castro, L., Rodríguez-Fernandez, N., Torrente-Patiño, Á., Carballal, A. *Artificial Neural Networks and Deep Learning in the Visual Arts: a review. Neural Comput & Applic* (2021) <https://doi.org/10.1007/s00521-020-05565-4>
- 2 Carballal, A., Cedron, F., Santos, I. *et al. Minimal neural network topology optimization for aesthetic classification. Neural Comput & Applic* (2021). <https://doi.org/10.1007/s00521-020-05550-x>

Autora: **Iria María Santos López**

Titor: **Juan Jesús Romero Cardalda**

Directores: **Adrián Carballal Mato** e **Juan Jesús Romero Cardalda**

Universidade da Coruña

- 3 Carballal, A., Fernandez-Lozano, C., Rodriguez-Fernandez, N., Santos, I., & Romero, J. (2020). *Comparison of Outlier-Tolerant Models for Measuring Visual Complexity*. *Entropy*, 22(4), 488.

A realización foi parella, aínda que as datas de publicación dependeron do proceso de edición das distintas revistas. Concretamente, primeiro publicouse o de **Entropy** e posteriormente os de **Neural Computing and Applications**. En cambio, a orde en que se fará mención nesta tese será a contraria á de publicación, posto que a coherencia narrativa e os traballos realizados foron levados a cabo cunha metodoloxía distinta.

Comezaremos a tese cunha introdución que nos servirá para incorporarnos aos distintos conceptos máis relevantes ao redor das diferentes artes, posto que é fundamental comprender o concepto de *arte* e cales son os seus tipos para poder afondar nos estudos relativos ao tema. Tamén dentro desta introdución abordaremos os distintos usos que se fai da **IA** aplicada ás artes.

Tras a introdución, inclúese un pequeno apartado de xustificación e obxectivos, para ter unha idea clara dos resultados esperados ao longo dos diferentes artigos.

A tese continúa cun resumo do contido dos distintos artigos cos seus datos máis relevantes, onde podemos ver tanto a discusión coma os resultados. O artigo 1 trátase dun estado da arte sobre o uso de redes de neuronas artificiais e *deep learning* nos distintos campos das artes visuais. O artigo 2 trátase dunha identificación da topoloxía mínima de rede neuronal para realizar unha categorización estética de imaxes. O artigo 3 céntrase en buscar, entre os distintos modelos de *machine learning*, o máis axeitado para obter resultados óptimos en relación coa complexidade visual, valor moi relacionado con preferencia estética en diversos traballos de psicoloxía da estética.

Os tres traballos son complementarios e afondan no uso de técnicas de *soft computing* para a predición de gusto estético ben directamente (artigo 2) ben a través da complexidade visual percibida polo usuario (artigo 3). O artigo 1 proporciona o estado da arte que serve de referencia inicial a esta investigación, proporcionando ademais de traballos relacionados, un conxunto de *datasets* de referencia co que se validan as propostas realizadas.

Autora: **Iria María Santos López**

Titor: **Juan Jesús Romero Cardalda**

Directores: **Adrián Carballal Mato** e **Juan Jesús Romero Cardalda**

Universidade da Coruña

Ao final desta tese presentaranse unhas conclusións globais. Por último, a modo de anexo, engádense tamén os tres artigos na súa totalidade no idioma en que foron publicados (inglés).

1.1 Belas artes e tarefas artísticas

Resulta importante para comezar esta tese afondar nas orixes da arte e en que podemos considerar como tal. Por exemplo, **Zeitlin e Winkler** (1990) consideraban erróneo chamar pezas de teatro no sentido moderno da palabra ao teatro grego. Na época dos antigos gregos, de feito, non tiñan unha palabra para denominar estas “artes”, senón que utilizaban *techné*, incluíndo tamén na definición do termo outras tarefas, que na actualidade se denominan “oficio”. O mesmo pasaba coas *ars* romanas (**Ritcher**, 1959; **Shiner**, 2004).

No medievo artista e artesán eran tratados coma un todo, de modo que o seu labor era un conxunto de habilidade, imaxinación, tradición e invención, e non foi ata o século XVII cando se produciu a transición cara ao que hoxe se coñece como “arte” (**Shiner**, 2004).

A continuación, veremos que existen varias disciplinas artísticas empregadas para agrupar aos diferentes tipos de arte na actualidade.

1.1.1 Disciplinas artísticas

As distintas disciplinas artísticas poden englobarse en **belas artes**, que crean unha impresión aos sentidos superiores (vista e oído), e **artes menores** que, neste caso, resultan do gozo de sentidos menores (gusto, olfato e tacto). As academias italianas (s. XVI) e a academia francesa (s. XVII) elaboraron un completo sistema formativo e profesional que foi modelo para todas as academias de Belas Artes de Europa e América a partir do século XVIII (**Prevsner**, 1982).

No século XVIII foi popularizado o termo “belas artes” e o primeiro libro coñecido por clasificalas foi o escrito por **Batteaux** (1747), *Les Beaux-Arts réduits à un même principe*. **Batteaux** (1747) quixo realizar unha unificación das teorías sobre beleza e gusto existentes durante a época. Incluíu dentro desta categoría a danza, a escultura, a música, a pintura e a poesía. Tamén considerou a elocuencia unha arte, posteriormente. Neste sentido, co paso dos anos, houbo cambios circunstanciais realizados por outros autores, que engadían ou quitaban “artes” á listaxe e foi **Canudo** (1911) o primeiro en denominar sétima arte o cinema. A oitava arte coñecida, por tanto, sería a fotografía, que mostra imaxes que poden afectar ao goce visual. Por

Autora: **Iria María Santos López**

Titor: **Juan Jesús Romero Cardalda**

Directores: **Adrián Carballal Mato** e **Juan Jesús Romero Cardalda**

Universidade da Coruña

outro lado, o cómic ten as súas orixes nas viñetas de **Rodolphe Töpffer** (profesor suízo), que comezou a combinar os seus debuxos con texto (**Revert Gomis**, 2016); e actualmente é considerado polos expertos a novena arte (**Lacassin**, 1971; **Christiansen**, 2000). Ademais, tamén poden considerarse **belas artes** os videoxogos, xa que comparten metas expresivas con outras formas de arte recoñecidas e, incluso, museos e programas de estudos dedicados á arte incorporan exhibicións e programas de estudos de videoxogos (**Smuts**, 2005). Por tanto, dentro das **belas artes** podemos distinguir varias disciplinas categóricas, pero que non teñen límites estritos nin excluíntes unhas das outras: artes visuais, artes escénicas ou de *performance*, artes musicais ou sonoras, artes literarias, artes aplicadas, artes plásticas, artes gráficas e artes baseadas no tempo. Neste sentido unha arte visual podería ser tamén arte aplicada, se se tratase, por exemplo, dunha fotografía xornalística. Pola súa banda, as **artes menores** son a gastronomía, a perfumaría ou a artesanía.

1.1.1.1 Artes visuais

A escola de arte alemá, **Bauhaus** (1919-1933) establece que as artes visuais son esencialmente unha linguaxe, cun vocabulario composto por elementos gráficos abstractos e básicos: liña, cor, punto, etc., con regras sintácticas que relacionan dous ou máis destes elementos entre eles. A definición é moi limitada, sendo específica da arte abstracta, pero pode estenderse a outros tipos de arte. O sistema pode levarse a cabo por medio de distintos materiais ou procesos, de modo que calquera obra profesional de debuxo, pintura, escultura, arquitectura, deseño gráfico ou industrial forma parte desta disciplina artística (**Viadel**, 1997). Seguindo a teoría, tamén podemos aglutinar nas artes visuais os videoxogos, o cinema, a fotografía e o cómic.

1.1.1.2 Artes escénicas ou de *performance*

As artes escénicas son as manifestacións humanas representadas por un ou máis suxeitos (bailaríns, autor, compañía de actores, etc.) nun espazo escénico, en tempo determinado e de cara ao público.

É habitual confundir espazo escénico con escenario teatral, pero na actualidade xa non é unha realidade única, senón que é posible realizar artes escénicas en espazos arquitectónicos ou urbanísticos construídos ou habilitados para a dita finalidade.

Podemos englobar dentro das distintas artes escénicas o teatro, a danza, a poesía declarada, a música e calquera outro espectáculo (**Cuenca et al.**, 2010), que cumpra cos requisitos mencionados con anterioridade.

Autora: **Iria María Santos López**

Titor: **Juan Jesús Romero Cardalda**

Directores: **Adrián Carballal Mato** e **Juan Jesús Romero Cardalda**

Universidade da Coruña

1.1.1.3 Artes musicais ou sonoras

Pode considerarse mediados do século XVI o momento en que a música comezou a tomar forma do que hoxe consideramos “arte” e, especificamente, a invención da ópera presenta o caso máis claro de música como arte (**Kivy**, 1991).

Outros autores como **Iges et al.** (2016) consideran os últimos decenios do século XIX a apertura do mundo occidental a este tipo de arte. En concreto 1980 é o ano en que os usos artísticos do son por parte de artistas, vídeoartistas, compositores, poetas experimentais e *performers* foron etiquetados baixo o termo “arte sonora”, aínda que houbo iniciativas que comezaron antes a empregar o son como elemento expresivo, por exemplo o cinema experimental.

Considéranse neste caso as manipulacións de sons e silencios de forma coherente co uso da melodía, o ritmo e a harmonía para compoñer pezas, que crean unha sensación na escoita, estimulan e producen distintas emocións a través de criterios artísticos.

No se trataría, por tanto, de dar una definición del término “arte sonoro” para categorizarlo o delimitarlo, sino de intentar pensar el sonido, de manera inédita, desde un ámbito que se encuentre alejado del privilegio de uno de los registros sensibles (ya sea el del oído o el de la mirada). Las preguntas que nos quedan por reconsiderar no parecen ser cronológicas, más bien se relacionan con el tipo de lugar que se abre con la escucha; con las formas en las que el sonido resuena en el espacio (y remite, a su vez, en nuestros cuerpos); con la posibilidad de deslizarse desde una fenomenología de la percepción (de la mera constatación, interactiva o no, de las propiedades acústicas del sonido) a una fenomenología de lo sensible, o, al fin, con el concepto de escucha que estaría en juego cuando se pretende dar cuenta del sonido en el arte.

José Luis Maire (Iges et al., 2016)

De aí que autores como **Lander** (1990) formulen discusión sobre a diferenza entre “arte sonora” e “arte musical”, pero que se resolve coa definición que **Cage** (1991) lle outorga á música, ata poder englobarse dentro dela todo o que pode ser escoitado. Así, unha música experimental empregada nunha obra de arte sonora sería considerada tamén música, polo tanto, se xuntamos as artes musicais e as sonoras

Autora: **Iria María Santos López**

Titor: **Juan Jesús Romero Cardalda**

Directores: **Adrián Carballal Mato** e **Juan Jesús Romero Cardalda**

Universidade da Coruña

podemos abarcar a disciplina que englobe todas as tarefas que conteñan música (**Dokic**, 1998).

1.1.1.4 Artes aplicadas

Calquera tipo de arte, xa sexa dentro das **belas artes** xa sexa das **artes menores**, cuxa aplicación é funcional, industrial ou decorativa, pero onde a estética ou plástica ten que ver directamente coa súa función considérase arte aplicada. Algúns exemplos deste tipo de artes son a fotografía xornalística, que xa mencionamos con anterioridade, a arquitectura funcional, a artesanía (especialmente a cerámica), a escultura, as artes gráficas (caso do deseño gráfico e da ilustración), o deseño industrial, a perfumaría, a xoiaría, a gastronomía, o deseño de interiores ou deseño téxtil (**Letelier**, 1937; **Feduchi**, 1946; **Alcolea**, 1975; **Maltese**, 1980; **Correa**, 1982; **Thomas et al.**, 1985; **Fleming e Honour**, 1987; **Thomsen Marcial**, 1995; **Gruber**, 2000a/b; **Rovira**, 2015).

1.1.1.5 Artes plásticas

Resulta complexo distinguir entre artes visuais e artes plásticas, porque todas as obras que pertencen a esta última categoría forman parte sempre das artes visuais, máis non necesariamente unha arte visual é plástica. Para considerarse arte plástica o autor ou autora debe utilizar a materia con que crea a obra de forma activa. Entendemos entón que dentro das artes plásticas se atoparían a arquitectura, a escultura, a pintura, o debuxo, o cómic e o gravado. Pola contra, non se considerarían dentro desta categoría os videoxogos, a fotografía, o cinema, nin o deseño gráfico ou industrial (**Souriau**, 1949).

En controversia, existen autores que consideran a imaxe fotográfica como unha arte plástica (**Greimas et al.**, 1989; **Bazin**, 2004), dende un punto de vista enfocado á fotografía analóxica, onde o revelado pode ser manual.

1.1.1.6 Artes gráficas

Pola súa banda, as artes gráficas, que naceron coa creación da imprenta de **Gutenberg** (**Fray et al.**, 2020), son todas as técnicas artísticas, que contan con impresión e que funcionan en dúas dimensións. Inclúense dentro desta disciplina tamén o cómic e o gravado, así como a litografía, a caligrafía (manual ou dixital) e as artes que contan con impresión dixital (deseño gráfico e fotografía).

1.1.1.7 Artes literarias

Partindo da caligrafía, temos tamén as artes literarias, que albergan as obras ao redor da palabra escrita, lida ou escoitada. Atópanse nesta categoría todas aquelas

Autora: **Iria María Santos López**

Titor: **Juan Jesús Romero Cardalda**

Directores: **Adrián Carballal Mato** e **Juan Jesús Romero Cardalda**

Universidade da Coruña

tradicions orais das distintas culturas, transmitidas ao longo das xeracions aínda que non estean recollidas de forma escrita. Son as creacions literarias que posúen mérito artístico e non se trata de palabras escritas de forma ordinaria. Tamén poden recollerse dentro da disciplina das artes literarias as novelas, a poesía, o guiión teatral ou cinematográfico, e o ensaio. (Parra Jaramillo, 2018)

1.1.2 Valor dos traballos artísticos

A palabra estética provén do termo grego *aesthesis*, que significa ‘sentimento ou sensibilidade’. Actualmente, o termo é interpretado como unha disciplina filosófica que estuda as condicións do belo na arte e na natureza (Wadjiz, 1980).

O valor dunha obra de arte defínese en moitos casos polo seu prezo económico; mais este prezo está directamente relacionado co pracer do espectador ou usuario sobre a dita obra de arte. Na actualidade, estamos rodeados de arte. Incluso existe unha rede social (**Instagram**), que se basea directamente na arte da fotografía. Ademais, resulta importante poder medir o gusto estético das persoas, porque en plataformas coma a citada as imaxes son compartidas regularmente a través de repeticións visuais e de xénero, que están conectadas directamente con xéneros visuais vistos en revistas e libros. Tales padróns visuais funcionan coma xerme do atractivo afectivo das imaxes dentro da cultura dixital visual (Leaver *et al.*, 2020). Como afirmou Mapes (2021), a estilización é un recurso persuasivo para a proliferación de valores, polo que medir o valor estético resulta esencial á hora de levar a cabo tarefas comerciais de *márketing* e publicidade, entre outras. Estas medidas foron a raíz de moitas investigacións científicas (McWhinnie, 1968; den Heijer e Eiben, 2010).

Existe constancia dos grandes esforzos xerais para formular unha psicoloxía da arte, que busca agrupar as características creativas, que resultan agradables para a percepción tanto de artistas coma do resto da poboación (Barron, 1963; Valentine, 2005) e clasificalas a través de medidas estéticas.

Dende 1933, diferentes investigadores en psicoloxía da estética buscaron factores que influencian o valor estético das imaxes. Birkhoff (2013) propón a primeira teoría da estética, na que a estética é proporcional á orde da imaxe e inversamente proporcional á súa complexidade visual (fórmula 1):

$$M = \frac{O}{C} \quad (1)$$

Eysenck (1942) propuxo un enfoque experimental para a teoría de estética da **Gestalt**, na cal a clave estética estaba na simplicidade, a regularidade e a simetría. A súa teoría dicía que a beleza era proporcional á diminución de cantidade de enerxía necesaria no acto de percibir o obxecto estético. A partir desta teoría creou dúas leis, que a completaban:

- A lei da retentividade: para completar e satisfacer, a forma esixe repetición.
- A lei da fatiga: para completar e satisfacer, o obxecto estético necesita variedade.

Woods e **Boudreau** (1950) compararon os hábitos visuais de artistas e non artistas a través dunha cámara bidimensional na cal estudaban a fixación ocular. Nestes experimentos descubriron, que, ao contrario do que falaba **Eysenck** (1942), no caso dos artistas a preferencia recaía sobre as áreas máis complexas.

Con base nestes traballos creouse unha liña de investigación en psicoloxía da estética que hoxe en día presenta publicacións especializadas como o *Psychology of Aesthetics, Creativity, and the Arts* e sociedades internacionais como a *International Association of Empirical Aesthetics*. Para unha análise máis profunda de traballos de psicoloxía da estética relacionados con predición de estética véxase **Johnson et al.** (2019).

1.2 IA, Machine Learning e Deep Learning

É importante distinguir entre **intelixencia artificial (IA)**, *machine learning* e *deep learning*, xa que a pesar de estares relacionados, tamén contan con diferenzas significativas.

A maioría das **IA** existentes son computadores creados para a resolución de problemas ou para a predición baseada en datos previamente introducidos. Utiliza os datos que ten para saber o que sucederá a continuación. Deste xeito, a **IA** sería a forma xenérica e dentro dela poden distinguirse tamén *machine learning* e *deep learning* (véxase Figura 1).

Autora: **Iria María Santos López**

Titor: **Juan Jesús Romero Cardalda**

Directores: **Adrián Carballal Mato** e **Juan Jesús Romero Cardalda**

Universidade da Coruña

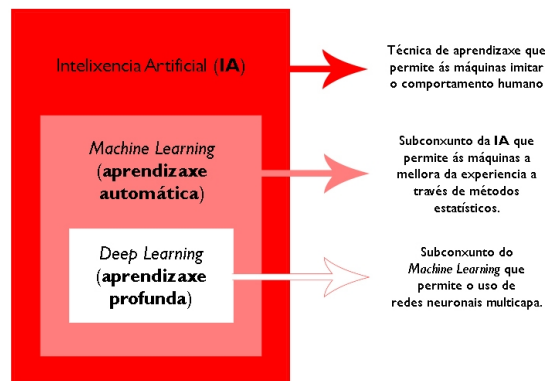


Figura 1: Gráfica que mostra a diferenza entre intelixencia artificial, machine learning e deep learning

Co *machine learning* (no que se inclúen as redes de neuronas) o ser humano ensínalle á máquina a recoñecer padróns baseados en datos e a realizar as predicións. Se prescindimos do *machine learning* (tamén coñecido como *aprendizaxe automática*) a IA estaría moi limitada e sería tan só unha programación sinxela como a seguinte:

$X = \text{verdadero} \rightarrow \text{acción } a$

$X = \text{falso} \rightarrow \text{acción } b$

Este era o uso do programa de xadrez de **Turing** (1988), por exemplo, onde se lle daban unhas determinadas características á máquina. Podemos poñer como exemplo a Figura 2, onde se mostra un saque realizado pola máquina. Neste exemplo “X” sería igual ao primeiro movemento dun peón. Se “X” é verdadeiro, por tanto, o peón non se moveu anteriormente, podería moverse dous cadros cara a adiante, coma neste caso, de e2 a e4. Do contrario, se “X” fose falso, e ao non se tratar do saque, polo xeral o único movemento posible sería moverse un cadro cara a adiante; exceptuando no caso no que houberse unha ficha ao seu redor para comer, no cal podería moverse un cadro en diagonal. Para isto, habería que ensinarlle á máquina paso a paso cada un dos movementos posibles.

Autora: Iria María Santos López

Titor: Juan Jesús Romero Cardalda

Directores: Adrián Carballal Mato e Juan Jesús Romero Cardalda

Universidade da Coruña

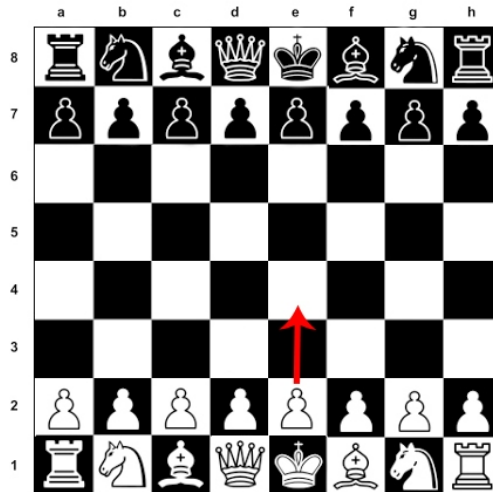


Figura 2. Movemento inicial peón no xadrez.

No caso do *machine learning*, que tivo lugar a partir dos anos oitanta do pasado século, dáselle á máquina un conxunto de datos para aprender, así como unhas características que determinan o conxunto, e posteriormente poderá identificar por si mesma similitudes en distintos contidos, sen ter que limitarse tan só a “X”.

Seguindo co exemplo do taboleiro de xadrez, podemos fixarnos na Figura 3. Como xa mencionamos anteriormente os movementos posibles do peón, exceptuando o saque, son dun cadro. Pode ser un cadro cara a adiante, polo xeral, ou no caso de comer figuras do contrincante, un cadro en diagonal. Neste caso, na figura móstrase que o peón situado en e4 tería dous movementos posibles. Podería moverse un cadro cara a adiante para situarse en e5 ou un cadro en diagonal para comer o peón negro situado en d5. Nunca poderá moverse cara a f5, xa que non hai ningunha figura do xogador contrario. A máquina deste xeito, tras terlle ensinado como características os distintos movementos do resto de figuras escollerá a mellor opción, que neste caso será comer o peón contrario, posto que non corre ningún perigo, dado que ningunha das figuras do contrincante pode acceder directamente ao cadro d5. Os únicos movementos posibles para o contrincante neste caso serían:

- Cabalo (b8): movementos en forma de L cara a d7, a6 ou c6.
- Alfil (c8): movemento en diagonal cara a d7, e6, f5, g4 ou h3.
- Rei (d8): un cadro cara a d7.
- Raíña (e8): movemento en diagonal cara a d7, c6, b5 ou a4.
- Cabalo (g8): movemento en forma de L cara a f6 ou h6.

Autora: Iria María Santos López

Titor: Juan Jesús Romero Cardalda

Directores: Adrián Carballal Mato e Juan Jesús Romero Cardalda

Universidade da Coruña

- Peóns: un ou dous cadros cara adiante situándose na fila 6 e 5, pero nunca 2 cadros en diagonal nin posibilidade de comer cara a adiante.

Todos estes e outros movementos tiveron que ensinárselle á máquina previamente, para que ela con posterioridade poida escoller o mellor movemento dentro das distintas posibilidades para cada figura.

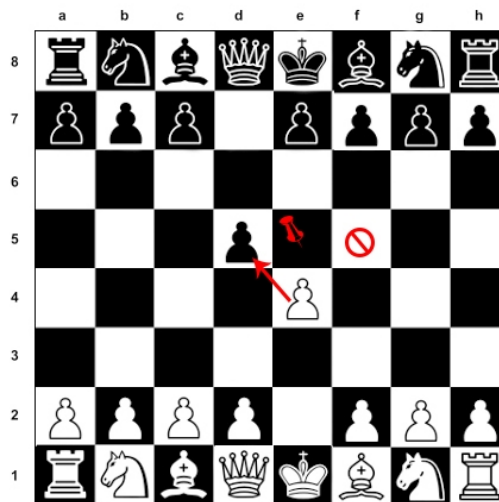


Figura 3. Movemento preferente de peón para comer ao peón do contrincante no taboleiro de xadrez.

Pola súa banda, a *deep learning*, introducida no ano 2010, trátase da **IA** onde a máquina é capaz de razoar e tirar as súas propias conclusións; aprende por si mesma. A *deep learning* é unha aprendizaxe automática con redes neuronais profundas, que procesan os datos de forma similar a como o fai o cerebro humano. Non é necesario darlle datos sobre a aprendizaxe que se lle quere dar á máquina, tan só mostrarlle un número amplo de resultados similares aos que ten que atopar, para que por si mesma poida realizar os cálculos necesarios e chegar a novos resultados similares aos dados. O exemplo utilizado con anterioridade do taboleiro de xadrez sería sinxelo tamén nese caso. Habería que mostrarlle á máquina varias partidas de xadrez (ou permitirlle xogar moitas partidas consigo mesma ou con outros) para que por si soa aprenda os movementos que pode levar a cabo e tamén cal é o método ideal para chegar ao xaque mate. Actualmente, a través deste método existen máquinas capaces de gañar a grandes profesionais de xogos de estratexia, dende o xadrez ata outros de maior complexidade como pode ser o Go¹.

1 Existe un documental titulado AlphaGo que trata sobre o uso de *Deep Learning* para vencer aos máis grandes xogadores de Go, que actualmente está dispoñible de forma gratuíta en Youtube no seguinte enlace: https://www.youtube.com/watch?v=WXuK6gekU1Y&ab_channel=DeepMind (04/03/2021).

1.3 Técnicas adaptativas

As técnicas adaptativas tratan de empregar mecanismos utilizados na natureza para resolver problemas computacionais. A través do uso deste tipo de técnicas búscase conseguir unha autonomía na aprendizaxe das máquinas: a programación automática. Nesta tese faremos un breve repaso das principais técnicas adaptativas, que sirva para comprender mellor o traballo posterior.

1.3.1 Redes de neuronas artificiais

As redes de neuronas artificiais ou sistemas conexionistas son un sistema inspirado no sistema nervioso biolóxico, que intentan simular o comportamento cerebral. Estas redes constan de elementos organizados en capas e conexións entre elas. Na Figura 4 móstrase de forma visual mediante un esquema o funcionamento das redes de neuronas artificiais. Os datos entran a diferentes capas de entrada e móvense por todas as capas ocultas ata chegar á resolución do problema, que é o que se presenta na capa de saída. *Deep learning* é o uso de modelos complexos de redes de neuronas artificiais.

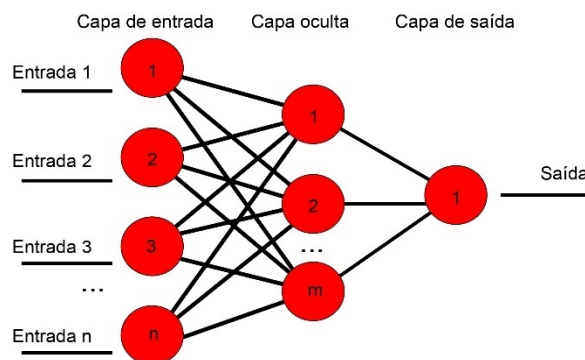


Figura 4. Funcionamento das redes de neuronas artificiais

1.3.2 Algoritmos xenéticos

Os algoritmos xenéticos, pola súa parte, son técnicas de computación evolutiva que se inspiran nos procesos de selección natural. Baséanse na teoría da evolución de **Darwin** en que os organismos que teñen a capacidade de adaptarse ao contorno para sobrevivir reproducense mentres que os que carecen dela morren (**Smith e Maynard**, 1993).

Os algoritmos xenéticos lévanse a cabo sobre poboacións de individuos que representan solucións ao problema a través de operacións de selección, cruzamento e mutación. Os individuos vanse combinando e a medida que avanzan as xeracións achéganse á resolución final (**Holland**, 1996). Na Figura 5 podemos ver de forma

gráfica este proceso, que se inicia coa poboación inicial, normalmente aleatoria. En cada xeración avalíanse os diferentes individuos segundo unha función de avaliación. O ADN dos individuos máis adaptados ten maior probabilidade de replicarse á seguinte xeración, que é obtida cos operadores de cruzamento e mutación.

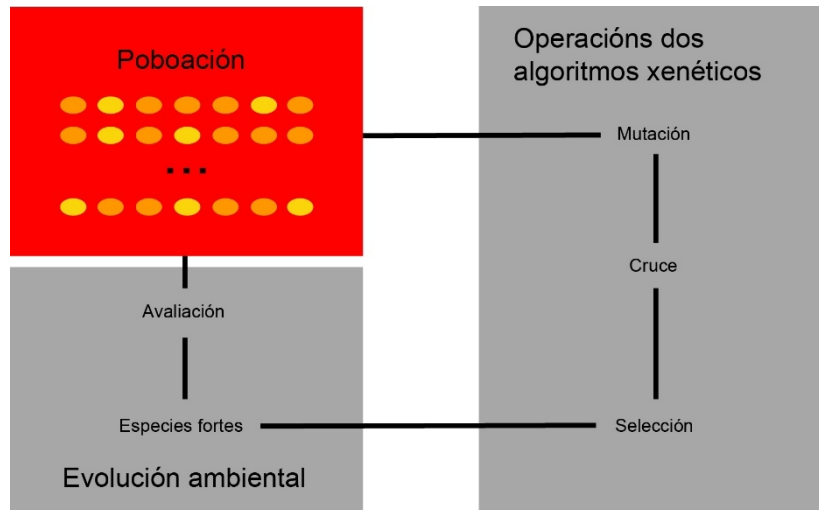


Figura 5. Funcionamento dos algoritmos xenéticos. O proceso comeza na poboación e vai avanzando en forma circular no sentido contrario ás agullas do reloxo.

1.3.3 Programación xenética

A programación xenética tamén forma parte da computación evolutiva e constitúe unha variante dos algoritmos xenéticos, que mantén o principio da selección natural, pero onde a través da combinación dun conxunto inicial de programas, que serve como poboación, se crean as distintas xeracións, que dan lugar a mellores programas (poboacións). A diferenza fundamental da programación xenética cos algoritmos xenéticos é que neste caso o sistema de codificación se realiza en forma de árbore. Podemos ver o proceso na Figura 6.

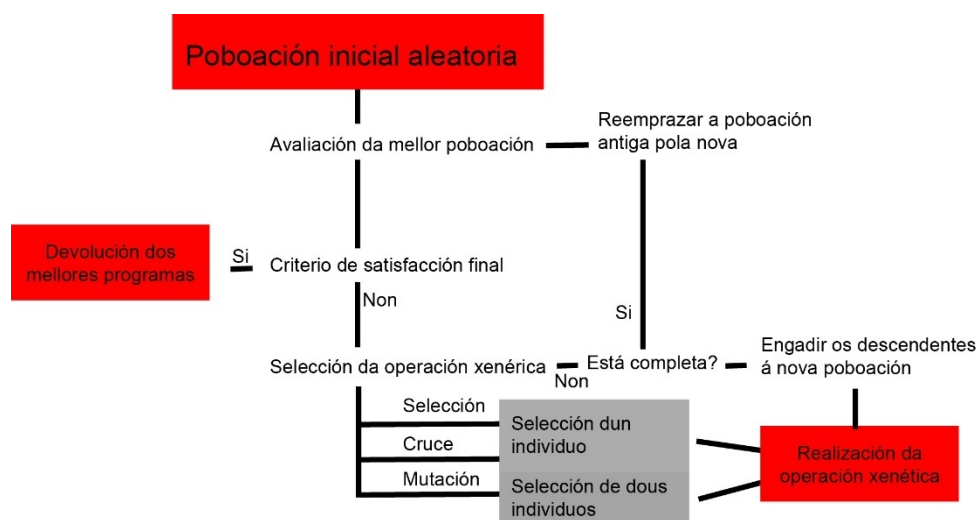


Figura 6. Funcionamento da programación xenética.

1.4 IA aplicada a tarefas artísticas

Arte e ciencia comparten unha relación longa e duradeira. Probablemente o exemplo máis recoñecido sexa o traballo de **Leonardo da Vinci**, que se inclinou por explorar esta relación. No século XIX ambos os campos tiveron unha ruptura, pero a transferencia cruzada entre os dominios seguiu existindo.

Existen numerosas investigacións, dende eses inicios ata a actualidade, relacionadas co uso da **IA** aplicada ás distintas artes: xeración de formas en pintura ou escultura (**Stiny e Gips**, 1971; **Gips**, 1975); deseño arquitectónico conceptual (**As et al.**, 2018; **Pena et al.**, 2021); outros exemplos en deseño (**Lansdown**, 1988; **Gero e Sudweeks**, 2012; **Gero**, 2014) como o deseño gráfico (**Wu**, 2020); realización de obras creativas (**Penousal e Santos**, 2007; **Machado et al.**, 2008; **Gero e Maher**, 2013; **McCormack e d'Inverno**, 2014; **López de Mántaras**, 2016; **Gülzow et al.**, 2020); análise das obras e o seu estilo (**Jafarpour et al.**, 2009; **Nadikattu**, 2014); clasificación do xénero das obras (**Zujovic et al.**, 2009; **Ahmed et al.**, 2019); detección de falsificacións (**Polatkan et al.**, 2009); copyright das obras (**White e Matulionyte**, 2019); estilismo de moda (**Wang e Rasheed**, 2014); creación de obras literarias (**Bringsjord e Ferrucci**, 1999; **Oliveira**, 2009; **Oliveira**, 2012; **Das e Ganbäck**, 2014; **Wang et al.**, 2016a); expresións faciais para animación 3D (**Wang et al.**, 2004); outros usos en videoxogos (**Mateas**, 2003; **Yannakakis e Togelius**, 2018; **Bory**, 2019), aplicación en fotografía (**Lukac**, 2017) ou música (**Roads**, 1980; **Bradford e Côté-Laurence**, 1995; **Dannenberg**, 2000; **De Mantaras e Arcos**, 2002; **Liu e Ting**, 2016); etc.

Autora: Iria María Santos López

Titor: Juan Jesús Romero Cardalda

Directores: Adrián Carballal Mato e Juan Jesús Romero Cardalda

Universidade da Coruña

O campo de investigación nomeado como “estética computacional” ten como obxectivo investigar a realización de tarefas estéticas de forma autónoma polo computador dun xeito similar ás realizadas polos humanos (**Neumann et al.**, 2005).

Cardoso e Romero (2007) preguntaban “¿Tiene sentido creer que un enfoque computacional de la creatividad nos ayudará a entenderla? ¿Merece la pena investigar herramientas y entornos computacionales que puedan contribuir a la creatividad humana? ¿Es factible construir programas que puedan considerarse creativos?” e ao longo do mesmo capítulo deixaban ver a resposta afirmativa para as tres cuestións. Pero, ten sentido crer que un enfoque computacional da estética nos axudará a entendela? Merece a pena investigar ferramentas e contornas computacionais que poidan contribuír a temas relacionados coa estética? É factible crear programas que acheguen contidos estéticos de relevancia?

A estética, ao igual que a creatividade, trátase dun campo de investigación complexo. A estética tamén conta con diversos factores, como vimos anteriormente, que complican a súa definición e valoración.

Deste xeito, débese estudar a estética computacional dende diversas perspectivas. No primeiro dos artigos que compoñen esta tese faise unha revisión do estado da arte das técnicas de creatividade computacional que empregan *deep learning* e redes de neuronas artificiais para tarefas artísticas. A división por seccións deste estado da arte foi a seguinte: detección, clasificación, avaliación, xeración e reconstrución. Estes son principalmente os usos que se fan da **IA** para as tarefas artísticas en todas as súas disciplinas e tipos. Dentro destas diversas perspectivas, as que poden incluír contidos ao redor da estética serían só a clasificación, a avaliación, a creación e a reconstrución ou mellora. Poderíamos clasificar as obras de arte en relación á súa estética, predicir o seu valor estético, crear obras estéticas de forma autónoma a través de programas de computador ou reconstruír e mellorar o contido das obras no que á estética respecta (**Li e Chen**, 2009; **Desnoyer e Wettergreen**, 2010; **Feng e Ting**, 2014; **Jin et al.**, 2016; **Malu et al.**, 2017).

Nesta tese imos enfocarnos no uso de técnicas adaptativas para a predición de preferencia de usuario para as artes visuais.

Autora: **Iria María Santos López**

Titor: **Juan Jesús Romero Cardalda**

Directores: **Adrián Carballal Mato e Juan Jesús Romero Cardalda**

Universidade da Coruña

2 Xustificación e obxectivos

Como observamos ata o momento, os seres humanos buscaron a capacidade de comprender a arte dende varios enfoques, dende o filosófico ao científico, dende o principio dos tempos. Co desenvolvemento da informática, pioneiros como **Ada Lovelace** ou **Alan Turing** consideraron a posibilidade de crear computadoras con habilidades creativas e artísticas.

Dende a perspectiva da **IA** e a *machine learning* existen varios traballos que propoñen métricas para resolver problemas relacionados coa estética, empregando xeralmente imaxes que proveñen de sitios *web* de fotografía, que inclúen as avaliacións dos usuarios. Un exemplo importante disto é o artigo publicado por **Datta et al.** (2006), que presenta un conxunto de datos con 3.581 imaxes do portal **Photo.net**, avaliadas polo seu valor estético e a orixinalidade, con dúas ou máis valoracións; logrando unha taxa de éxito do 68'77%. Co mesmo conxunto de datos, **Wong** e **Low** (2009) publicaron un modelo que logrou un 78'8%. Posteriormente, **Wang et al.** (2016b) chegaron a un 82'4% e unha correlación do 0.79, e **Tan et al.** (2016b) ao 86'67% cunha correlación do 0.843. Finalmente, **Xia et al.** (2017) lograron unha taxa de éxito do 86'14%.

Un artigo publicado polo **RNASA-IMEDIR** na *Neural Computing & Applications* (Q1) presentou unha correlación do 0'94 para predición e unha taxa de éxito do 93% na tarefa de clasificación (**Carballal et al.**, 2018). Existen traballos similares que empregan conxuntos de imaxes dos portais **Photo.net**, **DPChallenge.com** e **Flickr** (**Luo e Tang**, 2008; **Liu et al.**, 2010; **Romero et al.**, 2011; **Romero et al.**, 2012a).

Destes conxuntos de imaxes podemos atopar unha vantaxe (a dispoñibilidade de numerosas fotografías avaliadas por un gran número de usuarios) e diversos inconvenientes (**Romero et al.**, 2012b): as avaliacións dos usuarios poden depender dos gustos persoais, as relacións co creador do traballo, gustos particulares ou a moda do momento; non hai control sobre as condicións da realización da votación (tamaño das pantallas, distancia, iluminación, etc.); ademais o usuario pode avaliar factores non relacionados coa estética (**Datta et al.**, 2006) e pódese observar unha correlación alta entre a puntuación da “orixinalidade” e da “estética”; e as imaxes puntuadas con notas baixas teñen menos avaliacións.

Autora: **Iria María Santos López**

Titor: **Juan Jesús Romero Cardalda**

Directores: **Adrián Carballal Mato** e **Juan Jesús Romero Cardalda**

Universidade da Coruña

Hai poucos resultados na modelaxe estética, é dicir, na habilidade para realizar avaliacións e xuízos estéticos; pero o estudo foi abordado dende diversas perspectivas: **IA** (a través da estética computacional), a interacción humano-computadora (HCI), o *márketing* e a psicoloxía (a través da psicoloxía estética). Con respecto a HCI e *márketing* existen traballos que analizan a adaptación aos usuarios de vídeo (**Ben-Bassat et al.**, 2006; **Hartmann et al.**, 2008; **Lee e Koubek**, 2012; **Reinecke et al.**, 2013; **Bhandari et al.**, 2019) e sitios *web* (**Thielsch et al.**, 2019).

O obxectivo principal desta tese doutoral é validar a hipótese de que se pode predicir a preferencia estética media de usuarios mediante técnicas de **IA**.

Outros obxectivos son:

- Atopar un modelo neural óptimo, que permita a predición da preferencia estética media.
- Verificar a capacidade de métodos de redución de dimensionalidade para atopar conxuntos óptimos de métricas para predición estética.
- Analizar a capacidade computacional de predicir o valor de complexidade perceptual de imaxe e a súa relación coa predición de preferencia estética.
- Comparar diferentes métodos de *machine learning* en predición de complexidade perceptual.
- Analizar a achega actual de técnicas de *machine learning* e, en concreto, redes de neuronas artificiais e especialmente *deep learning*, a tarefas creativas e artísticas.

2.1 Metodoloxía

A metodoloxía é a adoptada polo grupo de investigación **RNASA-IMEDIR**:

- 1 Identificación do problema que hai que resolver
- 2 Xustificación do uso de técnicas de *soft computing* e *machine learning*
- 3 Deseño e construción dos conxuntos de datos
- 4 Adestramentos e ensaios
- 5 Organización e análise dos resultados
- 6 Validación dos sistemas
- 7 Repetición dos puntos 3, 4, 5 e 6 ata conseguir as aprendizaxes desexadas
- 8 Transferencia tecnolóxica

Autora: **Iria María Santos López**

Titor: **Juan Jesús Romero Cardalda**

Directores: **Adrián Carballal Mato** e **Juan Jesús Romero Cardalda**

Universidade da Coruña

2.2 Artigos

Inclúese nesta subsección unha tradución dos *abstracts* dos tres artigos científicos realizados e posteriormente, nos capítulos 3, 4 e 5, trataranse en profundidade.

2.2.1 Redes de neuronas artificiais e *deep learning* aplicada ás artes visuais: unha revisión

Neste artigo, realizamos unha análise exhaustiva do uso de redes de neuronas artificiais e aprendizaxe profunda nas artes visuais. Comezamos introducindo cambios na **IA** ao longo dos anos e examinamos en profundidade os últimos traballos realizados en predición, clasificación, avaliación, xeración e identificación a través de redes de neuronas artificiais para as distintas artes visuais. Aínda que destacamos as contribucións en fotografía e a arte pictórica, tamén hai outros usos como a modelaxe 3D, incluídos os videoxogos, a arquitectura e os cómics. Os resultados das investigacións discutidas mostran que o uso de redes de neuronas artificiais nas artes visuais segue evolucionando e experimentaron recentemente un incremento significativo. Para complementar o texto, inclúese tamén un glosario e unha táboa con información sobre os conxuntos de datos de imaxes máis empregados.

O capítulo 3 consta do contido referente a este artigo.

2.2.2 Optimización de topoloxía de rede neuronal mínima para clasificación estética

Neste artigo identifícase unha topoloxía de rede neuronal mínima en termos de unidades e conexións (neuronas e pesos), que permite avaliar estéticamente imaxes. Para este propósito, realizase unha análise para compactar a rede neuronal inicial reducindo a cantidade de conexións e neuronas necesarias para unha función de clasificación estética en capas ocultas. Utilizouse unha colección de fotografías e avaliacións estéticas dun sitio de votación en liña para fotografías (**Photo.net**). Estas imaxes están definidas por métricas *ad hoc* empregadas en estudos anteriores. Deste xeito, os datos de entrada, as imaxes, son interpretados dentro dun esquema de complexidade mínima por estimadores de complexidade (representacións internas). Os resultados obtidos mostran que a topoloxía optimizada atopada mantén o seu rendemento e a súa capacidade de xeneralización. Como parte deste estudo, atopouse unha relación sobre os datos de entrada e a topoloxía mínima necesaria

Autora: **Iria María Santos López**

Titor: **Juan Jesús Romero Cardalda**

Directores: **Adrián Carballal Mato** e **Juan Jesús Romero Cardalda**

Universidade da Coruña

para a súa correcta representación, demostrando estatisticamente un rendemento comparable a outras topoloxías non minimizadas.

No capítulo 4 afóndase no contido relativo a este experimento.

2.2.3 Comparación dos modelos máis tolerantes para medir a complexidade visual

Proporcionar a complexidade visual dunha imaxe en termos de impacto ou preferencia estética pode ser de grande aplicabilidade en áreas como a psicoloxía ou o *marketing*. Para este fin, certas áreas como “visión por computador” centráronse en identificar características e modelos computacionais que permitan obter resultados satisfactorios. Este traballo estuda a aplicación de modelos de *machine learning* recentes empregando imaxes de entrada avaliadas por humanos e caracterizadas polas particularidades relacionadas coa complexidade visual. Segundo os experimentos realizados, confirmouse que un destes métodos, “Correlation by Genetic Search” (CGS), baseado na busca de conxuntos mínimos de características que maximicen a correlación do modelo con respecto aos datos de entrada, predixeron valoracións humanas de complexidade visual da imaxe mellor que calquera outro modelo referenciado ata a data en termos de correlación, *RMSE* ou número mínimo de funcións requiridas polo modelo. Ademais, estudouse a variabilidade destes termos eliminando imaxes consideradas atípicas en estudos anteriores, observando a robusteza do método á hora de seleccionar as variables máis importantes para facer a predición.

O capítulo 5 explica a información sobre esta investigación.

Autora: Iria María Santos López

Titor: Juan Jesús Romero Cardalda

Directores: Adrián Carballal Mato e Juan Jesús Romero Cardalda

Universidade da Coruña

3 *Artificial Neural Networks and Deep Learning in the Visual Arts: a review*

Optouse por levar a cabo un estado da arte sobre o uso específico de redes de neuronas artificiais e *deep learning* nas artes visuais, posto que se trata dun campo do cal non existen estados da arte actualizados. Pola contra, si existen constantes avances no estado da arte para outros usos específicos da **IA** e doutras técnicas como a computación evolutiva. Por exemplo, **Lewis** (2008) publicou unha agrupación de 150 obras sobre computación evolutiva aplicada ás artes; **Galanter** (2012) fixo un obradoiro sobre avaliación estética computacional; **Spratt** e **Elgammal** (2014) analizaron o uso de visión por computadora na análise da pintura; **Toivonen** e **Gross** (2015) deron unha visión xeral dos posibles usos do *data mining* e o *machine learning* nos sistemas creativos; **Upadhyaya et al.** (2016) revisaron as funcións de alto e baixo nivel utilizadas para a recuperación de imaxes baseadas en contido (CBIR); **Johnson et al.** (2019) analizaron diferentes estudos sobre a psicoloxía da estética onde se inclúe a relación entre complexidade e estética, medidas de complexidade e preditores de complexidade dende a perspectiva da **IA**; e **Briot et al.** (2017) analizaron o uso de técnicas de *deep learning* aplicadas á xeración e análise de música.

Para a realización deste estado da arte seguiuuse unha metodoloxía de busca exhaustiva. Primeiro realizouse a busca en **Google Scholar** e **ResearchGate** de distintos traballos relacionados co tema, empregando palabras chave variadas, entre as que destacan: **AI (IA)**, *machine learning*, ANN (redes de neuronas artificiais), *image* (imaxe), *style transfer* (transferencia de estilo), *art generation* (xeración de arte), *artistic prediction* (predición artística), *visual art* (arte visual), *painting* (pintura), *painter* (pintor), e *aesthetics* (estética). Tamén foron consultados os artigos das distintas conferencias sobre **IA** aplicada a artes visuais e as publicacións especiais dos xornais, así como as conferencias **Bridges Conference**, **EvoMUSART** e **ICCC**. Posteriormente foron revisadas as citas dos artigos atopados. Finalmente filtráronse os resultados tendo en conta tan só nos que se fai emprego de técnicas de redes neuronais e desbotando outras técnicas. Ao longo do estado da arte faise referencia a determinados artigos, que fan emprego de *support vector machines (SVM)*, posto que

Autora: **Iria María Santos López**

Titor: **Juan Jesús Romero Cardalda**

Directores: **Adrián Carballal Mato** e **Juan Jesús Romero Cardalda**

Universidade da Coruña

presentan conxuntos de datos que foron utilizados noutros artigos nos que si se utilizan redes de neuronas artificiais.

Para a redacción, unha das fases máis complexas foi a de agrupar os artigos en temáticas similares, posto que se trata dunha tarefa bastante subxectiva e, incluso, existen artigos que poderían formar parte de diferentes seccións por conter varios dos contidos; pero fíxose unha criba o máis obxectiva posible, outorgándolles a temática de maior importancia dentro do traballo. As distintas seccións seleccionadas foron: detección de elementos en obras de arte; clasificación de acordo co autor ou estilo; clasificación baseada na calidade, complexidade e características visuais; avaliación da calidade ou estética; transferencia de estilo; e xeración ou reconstrución pictórica. Dentro de cada sección a orde dos artigos escollida foi a cronolóxica, para unha mellor mostra da evolución da tecnoloxía.

Ademais deste traballo, dentro do artigo fíxose tamén unha táboa resumo do número de cada unha das temáticas por ano, comezando no ano 2012, buscando deste xeito que o estado da arte fose o máis actual posíbel. Esta táboa foi tamén incluída a modo de índice nunha ligazón externa onde se pode acceder directamente aos artigos seleccionados.² Tamén se engadiu un anexo cunha táboa de conxuntos de datos en que se inclúe o acceso a todos os conxuntos dispoñibles publicamente empregados nos diferentes artigos.

A continuación, un resumo de cada unha das seccións mencionadas.

3.1 Detección de obxectos

Nesta sección foron incluídos os traballos dos últimos anos que empregaron redes de neuronas artificiais para a detección de elementos en pinturas e cómics, utilizando nalgúns casos conxuntos de datos de imaxe real para o adestramento.

Esta sección subdivídese en detección de obxectos en pinturas (5 artigos), detección de humanos en pinturas (1 artigo) e detección de obxectos en cómics (3 artigos), posto que se trata de tres usos diferenciados.

O primeiro traballo desta sección data do ano 2015 e nel logrouse superar o 40% na detección de obxectos a través de adestramento con fotografías e ensaio en pinturas. Dentro desta sección, os mellores resultados obtidos foron os de **Westlake et al.** (2016) a través de *Faster-CNN* que conseguiu unha precisión media do 58% tras axustar a rede neuronal convolucional nun conxunto de datos de debuxos animados e

² Pódese consultar o índice no seguinte enlace: <https://cutt.ly/4fCCWGs> (23/04/2021).

Autora: **Iria María Santos López**

Titor: **Juan Jesús Romero Cardalda**

Directores: **Adrián Carballal Mato** e **Juan Jesús Romero Cardalda**

Universidade da Coruña

imaxes de persoas. Tamén foron moi interesantes os resultados de **Dubray** e **Laubrock** (2019) para a segmentación automática de globos de texto cun *F1-score* ao redor do 0'94, onde F1 é a media harmónica entre e o recordo (**recall**), tamén nomeado como exhaustividade ou sensibilidade, e a **precisión** (fórmula 2):

$$F_1 = \frac{2 \times \text{precisión} \times \text{recall}}{\text{precisión} + \text{recall}} \quad (2)$$

Enténdese por **precisión** a identificación de proporcións positivas correcta (fórmula 3) e o **recall** coma a proporción de positivos reais identificados correctamente (fórmula 4).

$$\text{precisión} = \frac{VP}{VP + FP} \quad (3)$$

$$\text{recall} = \frac{VP}{VP + FN} \quad (4)$$

Nas fórmulas de **precisión** (3) e **recall** (4) VP equivale aos verdadeiros positivos, FP aos falsos positivos e FN aos falsos negativos.

O último dos experimentos desta sección é o de **Gonthier et al.** (2020), que tras ter realizados experimentos previos, utilizaron *Faster R-CNN* para a detección de obxectos sen supervisión e probárono en seis conxuntos de imaxes non fotográficas o que os levou a detectar varios problemas:

- Nalgunhas imaxes unha parte específica dun obxecto é detectada como completa.
- Algúns conxuntos son agrupados en instancias individuais de forma separada.
- O bordo da caixa ás veces é recortada de xeito incorrecto.

Autora: **Iria María Santos López**

Titor: **Juan Jesús Romero Cardalda**

Directores: **Adrián Carballal Mato** e **Juan Jesús Romero Cardalda**

Universidade da Coruña

- Existen imaxes cuxa complexidade semántica que crea fallos nas etiquetas.

Estes problemas son os que explican non ter logrado unha mellora en relación cos resultados anteriores dos outros autores.

3.2 Clasificación segundo o autor/a ou estilo

Os traballos integrados dentro desta sección son os que extraen características das artes visuais e as agrupan dependendo do seu autor/a ou estilo. Esta tarefa é máis complexa que a da subsección anterior, e quizais ese sexa un dos motivos polos que hai máis traballos relacionados co ámbito nos últimos anos. Nesta sección a división escollida foi a clasificación de pinturas e fotografías (17 artigos), a clasificación de debuxos (2 artigos), a clasificación de cómics (2 artigos), e a clasificación de obras arquitectónicas (1 artigo); a primeira das seccións (clasificación de pinturas e fotografías) é a que conta con maior número de traballos recentes.

Dentro da clasificación de pinturas e fotografías o primeiro traballo incluído non conta co uso de redes de neuronas artificiais, pero ten presenza destacada ao longo do artigo, posto que o conxunto de datos resultante foi utilizado por varios autores para validar os seus experimentos. Trátase do conxunto de datos de **Murray et al.** (2012) coñecido como “large-scale database for visual aesthetic” (**AVA**). O conxunto de datos é importante na área, xa que conta con máis de 250.000 imaxes cunha ampla variedade de metadatos. Outro *dataset* que se inclúe e tampouco conta con redes de neuronas artificiais é o de **Mensink e Van Gemert** (2014), polo mesmo motivo. Eles empregaron *1-vs-Rest linear SVM* con 112.039 reproducións de obras de arte de 6.629 artistas exhibidos no **Rijksmuseum**.

Xa dentro dos experimentos con redes de neuronas artificiais e *deep learning* o primeiro exemplo co que se contou foi o de **Karayev et al.** (2013). que presentaron dous conxuntos de datos: o primeiro de 80.000 fotografías de **Flickr** anotadas con 20 estilos e o segundo que contén 85.000 pinturas etiquetadas dentro de 25 categorías de estilo ou xénero utilizando *multi-layer network* para a predición de estilo. Lograron abarcar precisións de entre 72% e 94% na clasificación das distintas clases. Tamén recolleron 80.000 imaxes de **Pinterest** que, empregando o seu sistema, clasificaron en pinturas e fotografías.

Tan et al. (2016a) utilizaron *DCNN* para a clasificación por estilos, xéneros e autores, cuxos mellores resultados foron para a clasificación do Ukiyo-e (86%). Tamén

Autora: **Iria María Santos López**

Titor: **Juan Jesús Romero Cardalda**

Directores: **Adrián Carballal Mato e Juan Jesús Romero Cardalda**

Universidade da Coruña

obtiveron interesantes resultados en canto ao xénero de retratos (81%) e paisaxes (86%). Outros autores que lograron bos resultados foron **Bianco et al.** (2017) coa *novel deep multibranch neural network* en relación coa clasificación de estilo cun 85% de precisión; **Lecoutre et al.** (2017) a través do uso de *deep residual neural network* na predición de estilo con valores de precisión por clase dende 89'75% en diante; **Mao et al.** (2017) co uso do seu propio modelo de clasificación con *CNN (DeepArt)* cuxa precisión foi do 80%; e **Hicsonmez et al.** (2017) que catalogaron ilustracións co uso de *CNN* segundo o autor con resultados superiores ao 80% de precisión. O último experimento relativo a esta subsección foi o de **Hua et al.** (2020), que empregaron a combinación de *CNN* e *Markov Random Fields (MRF)* para a clasificación de autores.

Os autores que traballaron na clasificación de debuxos foron **Elgammal et al.** (2018) e **Sandoval et al.** (2019). O primeiro deles obtivo unha precisión do 75'7% co uso de *CNN* e o segundo cunha media de precisión do 77'53% empregando *Inception-v3* (**Jia et al.**, 2019).

Young-Min (2018) foi quen introduciu a clasificación de páxinas de cómic para cinco estilos distintos de manga xaponés utilizando unha *CNN* similar a *AlexNet* (**Kong et al.**, 2016), obtivo unha media de precisión do 86%. O mesmo autor propuxo no seguinte ano a clasificación para oito estilos de artistas diferentes segundo as páxinas e tamén os paneis e logrou unha media *F1-score* do 84% para páxinas completas e do 50% para paneis (**Young-Min**, 2019).

Finalizando coa sección de clasificación temos a **Yoshimura et al.** (2019), que trataron a clasificación de obras arquitectónicas a través do uso de *DCNN* adestrado utilizando fotografías de traballos de 34 arquitectos e obtiveron unha precisión do 73%.

3.3 Clasificación baseada na calidade, complexidade e características visuais

Ademais da detección de obxectos, tamén se empregaron redes de neuronas artificiais para a detección, clasificación e comparación de características visuais: complexidade, calidade, presenza de obxectos comúns... Os exemplos máis recentes relacionados coas artes visuais foron de detección, clasificación e comparación de semellanzas estilísticas. A división en subseccións comprendeu as tres do título da sección: clasificación segundo a calidade (2 artigos), clasificación segundo a complexidade (1 artigo) e clasificación segundo as características visuais (7 artigos).

Autora: **Iria María Santos López**

Titor: **Juan Jesús Romero Cardalda**

Directores: **Adrián Carballal Mato** e **Juan Jesús Romero Cardalda**

Universidade da Coruña

En relación coa clasificación segundo a calidade os dous artigos seleccionados son de dous anos non demasiado próximos no tempo. No primeiro, **Tian et al.** (2015) propuxeron un modelo de extracción de características abstractas a través do adestramento masivo cunha *DCNN* para empregalo na clasificación en imaxes de alta e baixa calidade e cuns resultados que consideraron óptimos a pesar de que un 20% das imaxes estaban mal clasificadas. No segundo, **Wagner et al.** (2019) aplicaron *deep learning* para a clasificación segundo a calidade das imaxes, pero non superaron os mellores resultados do método previamente empregado para o mesmo uso, **KonCept512**.

Tampouco foi moi tratado o tema da clasificación segundo a complexidade co uso de redes de neuronas artificiais, tendo o único exemplo con **Machado et al.** (2015). Este traballo emprega un modelo de *backpropagation*, que parte de métricas *ad hoc*, para predicir a complexidade perceptual dun conxunto de 800 imaxes, empregadas en experimentos psicolóxicos. Conta con resultados na correlación da predición e a puntuación media atribuída por humanos de R_s de ata un 0'834. Analízase despois un traballo incluído nesta tese, que aborda este problema (**Carballal et al.**, 2020). Por iso cómpre distinguir entre R_s e R^2 ou *R-Squared*, xa que son coeficientes diferentes empregados para avaliar resultados en experimentos.

R_s (fórmula 5) é o coeficiente de **Spearman** (tamén nomeado *rho*), que avalía a relación entre dúas variables continuas ou ordinais de forma monótona, é dicir, as variables tenden a cambiar ao mesmo tempo, pero non necesariamente a un ritmo constante. Pola súa banda, R^2 (*R-Squared*) é un coeficiente de determinación que equivale ao cadrado de r (o coeficiente de correlación **Pearson** (fórmula 6)), que avalía a relación das variables de forma lineal. No caso do coeficiente de Pearson, o cambio dunha variable asóciase ao cambio proporcional na outra. O valor de ambos os coeficientes pode variar de -1 a +1. Para que o valor sexa +1 en ambos os casos cando unha variable aumenta aumenta e a outra aumenta de forma consistente. Nos dous casos o coeficiente será cero se a correlación é case inexistente e, tamén para ambos os coeficientes, se a relación é decrecente o valor será -1. Nos casos nos que se diferencian son os seguintes: se unha variable aumenta e a cantidade non é consistente será menor de +1 en relación con **Pearson**, pero igualmente positivo, en cambio, no coeficiente de **Spearman** seguirá sendo +1; e se unha variable diminúe cando a outra aumenta, pero a cantidade non é consistente, entón o coeficiente de **Pearson** será negativo, pero maior que -1, mais no coeficiente de **Spearman** seguirá

Autora: **Iria María Santos López**

Titor: **Juan Jesús Romero Cardalda**

Directores: **Adrián Carballal Mato** e **Juan Jesús Romero Cardalda**

Universidade da Coruña

sendo -1 (**Restrepo** e **González**, 2007; **Chok**, 2010; **Hauke** e **Kossowski**, 2011; **Kasuya**, 2019).

$$r_s = \frac{\sum_{i=1}^n ((\text{rank}(x_i) - \overline{\text{rank}(x)}) (\text{rank}(y_i) - \overline{\text{rank}(y)}))}{\sqrt{\sum_{i=1}^n (\text{rank}(x_i) - \overline{\text{rank}(x)})^2 \sum_{i=1}^n (\text{rank}(y_i) - \overline{\text{rank}(y)})^2}} \quad (5)$$

$$r = \frac{\sum_{i=1}^n ((x_i - \bar{x})(y_i - \bar{y}))}{\sqrt{\sum_{i=1}^n (x_i - \bar{x})^2 \sum_{i=1}^n (y_i - \bar{y})^2}} \quad (6)$$

Finalmente, e cun número máis amplo de traballos, temos a clasificación segundo as características visuais, cuxo primeiro exemplo data do ano 2016 e que foi progresando paulatinamente ano tras ano, tendo o último artigo data do 2020.

Os resultados máis destacables foron os de **Carballal et al.** (2018), que utilizaron unha rede de neuronas artificial para distinguir entre pinturas e fotografías, obtendo para ambas as categorías resultados superiores ao 94% de taxa de coincidencia; e os de **Prasad et al.** (2018), que propuxo o uso de *CNN* para a clasificación de imaxes de flores cunha precisión máxima do 97'78%.

É interesante mencionar tamén o último dos traballos, no que **Castellano** e **Vessio** (2020) presentaron unha ferramenta na que empregaban *DCNN* para extraer características visuais de pinturas dixitalizadas para buscar similitudes entre estilos, mostrando deste modo posibles influencias. Algunhas das similitudes atopadas foron en pinturas de **William Turner** con **Rembrandt**, **Salvador Dalí** e **Gustave Courbet**; pinturas de **Alfred Sisley** con **Camille Pissarro** e **Claude Monet**; ou pinturas de **Sandro Botticelli** con **Michelangelo** e **Raffaello**.

3.4 Avaliación da calidade fotográfica e o valor estético

Esta sección inclúe dúas partes ben diferenciadas: a avaliación da calidade fotográfica (5 artigos) e a avaliación de valor estético (21 artigos). Trátase dunha

sección cuxos principios son bastante subxectivos, polo cal resulta significativo obter bos resultados relacionados con ambos os aspectos.

O primeiro que se tratou foron os conxuntos de datos máis relevantes empregados dentro de ambos os campos e posteriormente incluíronse os artigos dentro de cada subsección. Os principais conxuntos de datos empregados son **Photo.net**, **DPChallenge.com** e **AVA**.

Os traballos recentes de relevancia sobre avaliación de calidade fotográfica os fixeron **Tan et al.** (2016b), **Gao et al.** (2017), **Meng et al.** (2018), **Talebi e Milanfar** (2018), e **Zhang et al.** (2018). Os mellores resultados foron obtidos por estes últimos, que lograron promedios de entre 0'457 e 0'968.

As redes de neuronas artificiais tamén se utilizan en avaliación estética, que, de feito, é a sección con máis traballos recentes. **Carballal et al.** (2014) xa lograban uns resultados superiores ao 85% utilizando redes neuronais para a tarefa de clasificación binaria de composición estética para imaxe e vídeo. **Jin et al.** (2016) lograron tamén bos resultados a través de *Inception* con precisións de 85'62% e 79'25%. Os seus resultados os melloraron **Tan et al.** (2017) cunha precisión do 87'1%. **Li et al.** (2017) propuxeron unha *embedded learning convolutional neural network* (ELCNN) para a avaliación estética e os seus resultados por cada categoría non baixaron do 93% e son os máis altos de todos os abordados.

3.5 Transferencia de estilo e xeración pictórica ou reconstrución

Nos últimos cinco anos moitos investigadores utilizaron redes de neuronas artificiais e *deep learning* para a tarefa da transferencia de estilo, que consiste en transferir o estilo dunha imaxe ou conxunto de imaxes a outra. Son sete os artigos dentro desta sección e os resultados poden verse en exemplos visuais como o da Figura 7.

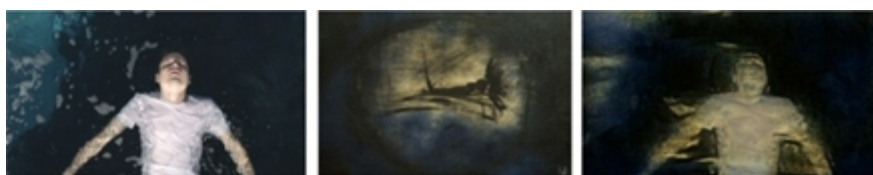


Figura 7. Resultados de Joshi et al. (2017) co uso de Neural Style Transfer co estilo impresionista de "Come Swim".

Autora: **Iria María Santos López**

Titor: **Juan Jesús Romero Cardalda**

Directores: **Adrián Carballal Mato** e **Juan Jesús Romero Cardalda**

Universidade da Coruña

Na Figura 7 podemos ver tres imaxes. A primeira imaxe é a imaxe orixinal (“Come Swim”), a segunda o estilo impresionista que se lle quere transferir e a terceira o resultado obtido (**Joshi et al.**, 2017).

A última sección dentro do artigo sobre o estado da arte do uso de redes de neuronas artificiais e *deep learning* en artes visuais abarca a tarefa máis complexa dentro da **IA** neste campo, xa que se trata da creación de imaxes. Nesta sección foron incluídas tamén reconstrucións de obras de arte, modelaxes 3D e creación de contidos gráficos para videoxogos. O número de artigos dentro desta sección foron 15 sobre xeración ou reconstrución fotográfica ou pictórica, 2 de xeración de modelos 3D e outros 2 referidos aos contidos gráficos dos videoxogos: un total de 19 artigos.

Nestes campos de xeración empréganse en moitos casos *deep convolutional generative adversarial networks (DCGANs)*. Unha **GAN** ten dúas redes de neuronas adestradas en paralelo: unha xera as imaxes segundo un criterio e a outra intenta identificar cales son as imaxes fóra do criterio.



Figura 8. "St George Killing the Dragon" feito por AICAN³ e vendido en Nova York (novembro do 2017) por 16.000\$ (Elgammal, 2019).

Os resultados para esta sección tamén poden verse mellor de forma visual, polo que se lles solicitou permiso aos autores para incluír as súas imaxes dentro do artigo. Pódense ver exemplos dos resultados na Figura 8, a Figura 9 e a Figura 10 dos resultados obtidos por **Elgammal** (2019), **Blair** (2019) e **Li et al.** (2019) sucesivamente.

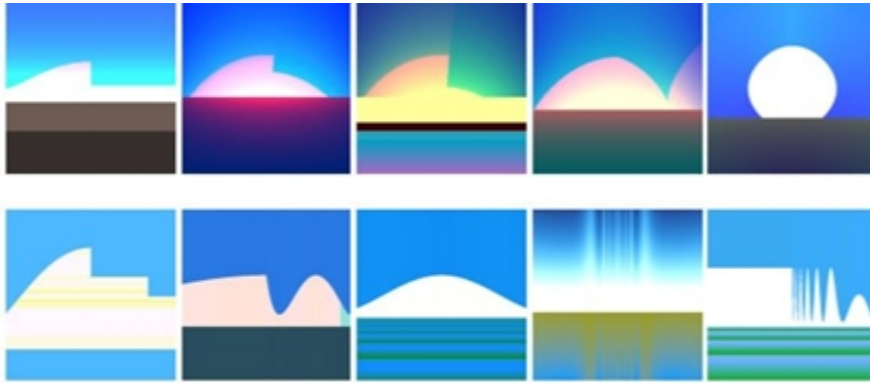


Figura 9. Exemplo dos resultados de xeración de imaxe con Hercule LeNet (Blair, 2019).

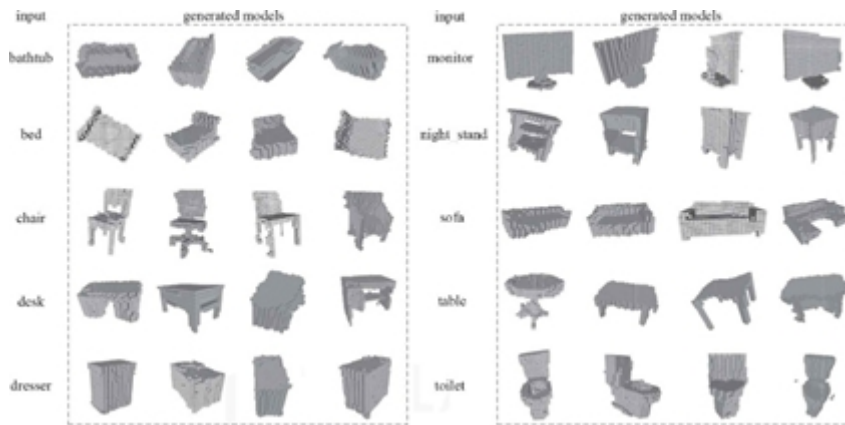


Figura 10. Exemplo de xeración de obxectos 3D nos experimentos de Li et al. (2019) cuxos resultados de precisión media no conxunto de datos IKEA foron do 70'9%.

Existen varios exemplos do uso de redes de neuronas artificiais e *deep learning* tanto para xeración como para transferencia de estilo, que non se presentan directamente en artigos de investigación, pero cuxos autores comparten a través da plataforma *online* **GitHub**⁴ como por exemplo **Surma** co seu **ADNet**⁵, que foi empregado nun dos artigos de investigación dos máis actuais (**Tian et al.**, 2020) sobre redución de ruído.

4 Algúns dos exemplos publicados en **GitHub**:

- https://github.com/HvyD/DeepLearning_Art_Generation
- <https://github.com/AhmetFurkanDEMIR/Deep-Learning-11-Art-generation-with-Neural-Style-Transfer>
- <https://github.com/utkarshj1303/Art-Generation-with-Neural-Style-Transfer-Coursera-Deep-Learning-Convolutional-Neural-Networks>
- https://github.com/maurock/neural_transfer_style
- <https://github.com/samiramunir/Classifying-Painting-Art-Style-With-Deep-Learning>
- <https://github.com/ashu10832/Style-Transfer-With-Deep-Learning>
- <https://github.com/wgcv/Style-Transfer-with-Deep-Learning>
- <https://github.com/Sarapuce/Transfert-Style>
- <https://github.com/aabidshamji/style-transfer>
- <https://github.com/umit-ozturk/Deep-Learning-Artistic-Style-Transfer-with-VGG16>
- <https://github.com/markerlich/GipStyleTransfer>
- <https://github.com/WSRogers98/Neural-Style-Transfer>
- https://github.com/sarthakbatragatech/style_transfer
- <https://github.com/stembarb/DL--Image-Style-Transfer>
- <https://github.com/nicktheodore/DeepLearningArt>
- <https://github.com/aseemarora1995/artistic-style-transfer>
- https://github.com/th-rpy/Cnn_Ann
- https://github.com/Caeri-Poot/Poot_TransferStyle_P
- <https://github.com/prathampatel9/Style-transfer>
- <https://github.com/chetanaramaiyer/styletransfer>
- https://github.com/Mirgahney/Nural_Art

(13/04/2021)

5 **ADNet** de **Surma** dispoñible en **GitHub**: <https://github.com/helloxiaotian/ADNet>
(13/04/2021)

Autora: **Iria María Santos López**

Titor: **Juan Jesús Romero Cardalda**

Directores: **Adrián Carballal Mato** e **Juan Jesús Romero Cardalda**

Universidade da Coruña

4 *Minimal Neural Network Topology Optimization for Aesthetic Classification*

Como se viu no estado da arte anterior, a determinación do valor estético é a sección máis traballada polos diferentes investigadores no campo nos últimos anos. Isto débese a que determinar o valor estético sería unha grande achega teórica e práctica en campos como a visión por computador e a psicoloxía. A relevancia da complexidade da imaxe percibida trátase dun tema recorrente no campo da estética. Isto é, que o valor estético está relacionado co pracer sensorial e intelectual, que aparece ao atopar unha percepción compacta dun estímulo visual complexo. Segundo **Machado e Cardoso** (1998) a relación entre estética e complexidade foi explorada en varios traballos de psicoloxía e gráficos por computadora e deu como resultado que a complexidade dunha imaxe está directamente relacionada coa entropía e a información mínima requirida para construír a imaxe, e inversamente relacionada coa orde. Deste xeito, pódese afirmar que depende do grao de previsión de cada píxel, polo que unha imaxe plana é menos complexa e unha imaxe aleatoria pura pódese ver como extremadamente complexa, de modo que é imposible a predición do valor de cada píxel, incluso en píxeles veciños. Pola súa parte **Liu e Chang** (1997) afirman que a complexidade depende de quen está a codificar ou percibir o obxecto. Este artigo enfócase na busca dun modelo que poida representar o xuízo estético atendendo á súa forma mínima compacta. O que é o mesmo, atopar un programa mínimo que permita predicir a estética sen perder rendemento nin capacidade ao compactalo.

As redes de neuronas son un conxunto de unidades (neuronas) conectadas entre si para transmitir sinais. A información de entrada pasa pola rede neuronal (onde sofre varias operacións) producindo valores de saída. Cada neurona está conectada a outras a través de ligazóns. Nestes enlaces, o valor de saída da neurona anterior é multiplicada por un valor de peso. Tales pesos poden aumentar ou paralizar o estado de activación das neuronas adxacentes. Igual que o cerebro humano, a rede aprende adaptando as súas conexións para crear unha representación interna dos datos de entrada, que en moitos casos non ten por que ser compacta; pode haber neuronas ou conexións que non interfiran na representación interna. Nestes casos poderían eliminarse da estrutura sen afectar á representación.

Autora: **Iria María Santos López**

Titor: **Juan Jesús Romero Cardalda**

Directores: **Adrián Carballal Mato e Juan Jesús Romero Cardalda**

Universidade da Coruña

Este artigo consiste en identificar un modelo de rede neuronal mínimo en termos de unidades e enlaces (neuronas e pesos), que permita clasificar as imaxes en termos de estética visual. Para iso, realizouse un estudo para compactar unha rede neuronal inicial minimizando o número de conexións e neuronas necesarias nas capas ocultas nunha tarefa de clasificación estética. O conxunto de imaxes e avaliacións empregado obtívose dunha plataforma de votación de imaxes en liña (**Photo.net**). Estas imaxes contan con métricas *ad hoc* empregadas en estudos anteriores. Así, os datos de entrada (imaxes) son representados por estimadores de complexidade (representacións mínimas) dentro dun sistema de complexidade compacto.

Tomouse a decisión de minimizar a topoloxía dunha rede completamente conectada, intentando minimizar o número de elementos por capa, xa que cada un dos elementos se conectará co resto das seguintes capas. Existiría a opción de minimizar as conexións e os elementos de procesamento implicando un aumento exponencial do tempo necesario para atopar posibles solucións. Outra forma de buscar o esquema de mínima complexidade sería facer un estudo das imaxes de entrada, ver cales delas ofrece información importante na aprendizaxe e cal non; mais non se levou a cabo deste xeito, xa que se pretende establecer algún tipo de regra para determinar o número mínimo de elementos necesarios na percepción multicapa (**MLP**) para garantir un bo rendemento e xeneralización, que sirva para determinar tipoloxías que minimicen o tempo de computación no adestramento.

4.1 *Dataset* e traballos previos

Datta et al. (2008) propuxo catro conxuntos de datos de probas para a realización de experimentos relacionados coa estética fotográfica. Todos os conxuntos de datos corresponden a portais de fotografía onde os usuarios poden facer unha avaliación de cada imaxe: **Photo.net**, **DPChallenge.com**, **Terragalleria.com** e **Alipr.com**. **Datta et al.** (2008) proporcionaron un arquivo que contén os identificadores de 20.278 imaxes e os seus votos dentro de cada categoría, que só contén imaxes de **Photo.net** (este conxunto está entre os recollidos en **Santos et al.**, 2021). Para o traballo foi posible recuperar aproximadamente 18.105 imaxes, posto que as restantes xa non están dispoñibles no sitio *web* ou a información recuperada non ten consistencia (os datos das puntuacións de cada categoría e das puntuacións medias ponderadas non coinciden).

Autora: **Iria María Santos López**

Titor: **Juan Jesús Romero Cardalda**

Directores: **Adrián Carballal Mato** e **Juan Jesús Romero Cardalda**

Universidade da Coruña

O conxunto de datos foi escollido por dous motivos: trátase do máis referenciado en clasificación e predición, e é o único que establece o número de votos para cada valor en cada imaxe. Os demais conxuntos tan só conteñen o valor estético medio. Neste caso sábese cantas persoas asignaron determinado valor a unha imaxe concreta.

Outro dos conxuntos máis empregados en estética é o de **Ke et al.** (2006), que recolle fotografías de **DPChallenge.com**. Foi un conxunto criticado por estar dividido en datos de adestramento e de probas. O terceiro conxunto a grande escala foi **AVA (Murray et al., 2012)**, que ofrece máis de 250.000 fotografías dun amplo número de temas e que inclúen metadatos entre os cales están as puntuacións estéticas de cada imaxe, pero este conxunto de datos non ten uns precedentes amplos de uso.

O conxunto de datos de **Datta et al.** (2008) utilizouse en experimentos de regresión e clasificación. Na regresión os mellores resultados de **Spearman** obtidos foron 0'5217 e 0'5826. Utilizando **AlexNet (Kong et al., 2016)** e **Inception v-3 (Jia et al., 2019)** sucesivamente. Ambas as redes foron adestradas previamente con **ImageNet**, un conxunto de datos composto por 15 millóns de imaxes de alta resolución e afinados coas imaxes de **Photo.net**. Os dous casos contan con só 16.666 imaxes das 20.278 definidas por **Datta et al.** (2008) e só 666 foron empregadas para validación.

En clasificación, os mellores resultados tiñan unha precisión do 65'37% a través do uso de **SVM** cunha validación cruzada de cinco pasos. Os datos que recollen por cada imaxe son os seguintes:

- Identificador de imaxe dentro do portal
- Número de votos
- Desviación media
- Desviación estándar
- Votos por cada categoría (valor mínimo: 1 – valor máximo: 7)

Resulta importante que en **Photo.net** as puntuacións están no rango do 1 ao 7, xa que noutros portais o intervalo vai do 1 ao 10, por exemplo en **DPChallenge.com**.

Grazas aos datos pódese facer unha análise que permita realizar modificacións para tratar de obter un conxunto de adestramento significativo. A idea é facer unha clasificación binaria formando dous conxuntos separados dos extremos da distribución. Foi unha metodoloxía aplicada en experimentos similares previamente.

Autora: **Iria María Santos López**

Titor: **Juan Jesús Romero Cardalda**

Directores: **Adrián Carballal Mato** e **Juan Jesús Romero Cardalda**

Universidade da Coruña

Datta et al. (2008), por exemplo, presentan a “fenda estética” onde fan a selección dos extremos da distribución na procura de dous subconxuntos suficientemente representativos establecendo tres grupos de imaxes diferenciáveis: baixa, media e alta calidade. Tras identificalos descartan o conxunto de calidade intermedio e realizan os seus experimentos de clasificación binaria cos restantes. Determinaron que as imaxes cun valor medio inferior a 4’2 irían ao conxunto de baixa calidade e as de valor medio superior a 5’8 aos de alta calidade.

Neste artigo realizouse un enfoque distinto para obter dous conxuntos heteroxéneos a través de dúas modificacións:

- Non se teñen en conta as imaxes que non foron valoradas por polo menos dez persoas distintas. → Estandarizar o número mínimo de avaliacións
- Calculouse a media e a desviación estándar das puntuacións e elimináronse nas que o consenso das puntuacións fose baixo. → Descartar imaxes con votos dispares

Tras ordenar o conxunto pola media calculada novamente, elíxense diferentes porcentaxes cada vez máis baixas para comprobar o funcionamento das redes xeradas con diferentes conxuntos de adestramento: metodoloxía empregada con anterioridade por **Ke et al.** (2006) para estudar o comportamento dun clasificador de imaxes profesional-afeccionado.

Decidiuse o uso de catro porcentaxes distintas para cada extremo: 5%, 10%, 15% e 20%. Deste xeito hai catro conxuntos de adestramento que van dende 1001 a 4004 imaxes con etiquetas de clasificación binaria proporcionada.

A elección destes “estratos” mediante porcentaxes pechadas permite a xeración de grupos diferentes sen atención ás imaxes. Podería utilizarse a información visual das imaxes, pero forzaría algún tipo de criterio para agrupar as imaxes dentro de cada subgrupo e podería implicar a aparición dalgún nesgo visual, coma o de contido, por exemplo. Hai que ter en conta a superposición real entre as puntuacións, isto é, puntuacións dentro de cada extremo que cae no lado oposto. Xa que existen casos marxinais nos que unha imaxe de baixa calidade obtivo unha nota distinta, por exemplo un 6; e o mesmo pode pasar cos extremos contrarios. Ao determinar as superposicións dos catro grupos de adestramento contouse coa referencia do erro intrínseco dos datos de entrada.

Autora: **Iria María Santos López**

Titor: **Juan Jesús Romero Cardalda**

Directores: **Adrián Carballal Mato** e **Juan Jesús Romero Cardalda**

Universidade da Coruña

Táboa 1: Valor de superposición e valores máximos e mínimos en cada unha das porcentaxes. As cores equivalen aos distintos adestramentos: laranxa (5%), azul (10%), verde (15%) e amarelo (20%).

Mínimo superior	Máximo inferior	Superposición (%)	Superposición baixa (%)	Superposición alta (%)	Precisión máxima estimada (%)
4'14	6'24	8'42	13'3	3'53	91'58
4'4	6'04	11'83	18'56	5'09	88'17
4'6	5'92	15'5	24'44	6'56	84'5
4'75	5'8	19'03	29'93	8'13	80'97

Os valores que poden visualizarse na Táboa 1 son:

- Mínimo superior: puntuación media mínima a partir da que se considera unha imaxe pertencente ao subconxunto de adestramento alto.
- Máximo inferior: puntuación media máxima a partir da cal se considera pertencente ao subconxunto baixo.
- Superposición: porcentaxe de votos baixos de imaxes altas e viceversa.
- Superposición baixa: porcentaxe de votos altos de imaxes baixas.
- Superposición alta: porcentaxe de votos baixos de imaxes altas.
- Precisión máxima estimada: estimación da precisión máxima do obxectivo que deben permitir estes datos de entrada.

A superposición no conxunto baixo sempre é maior, xa que nas imaxes altas hai máis similitude e nas baixas existe ampla diversidade. Por tanto, as valoracións persoais que se fan poden deberse ao contido, posto que non se trata de imaxes tan profesionais nin tratadas.

4.2 Métricas empregadas

Machado et al. (2015) suxeriron un gran número de posibles novas estimacións de complexidade baseadas na compresión de imaxes e a lei de **Zipf**. Unha vez que se analiza cada imaxe convértese de xeito independente nunha imaxe de tres canles **RGB** (vermello, verde e azul) cunha profundidade de 8 bits cos valores

Autora: **Iria María Santos López**

Titor: **Juan Jesús Romero Cardalda**

Directores: **Adrián Carballal Mato** e **Juan Jesús Romero Cardalda**

Universidade da Coruña

de píxeles nun intervalo [0:255] e cun tamaño predeterminado de 256x256. Así as imaxes enviadas serán do mesmo tamaño e forma.

A imaxe transfórmase logo á cor **HSV** (matiz, saturación e valor) procesándose así unha imaxe en escala de grises cunha canle para cada cor. A multiplicación píxel por píxel das canles **S** e **V**, e a multiplicación resultante a 255 tamén se empregan para xerar a nova imaxe en escala de grises. O novo método de cálculo **H** denomínase **HCS**, que é unha nova alternativa potencial ao número actual do modelo de cor de alto valor **HSV**.

Posteriormente realízanse tres transformacións, todas con **Canny** e **Sobel**, unha das cales aplica o filtro ao lado horizontal, a segunda á vertical e a terceira ás dúas. Calquera das transformacións engádese a cada imaxe orixinal e ás imaxes producidas eliminando as distintas canles de cor.

As desviacións media e estándar mídense coa intensidade do píxel de cada sinal, a excepción da imaxe da canle **H** (matiz). Estas desviacións determínanse dependendo da norma e do ángulo dos valores da tonalidade, debido a que a canle da tonalidade é circular. O valor do ángulo de tonalidade multiplícase polo valor **CS** e determínase unha norma para os valores de tonalidades e **CS** consecuentemente. Para as canles crecentes (**H**, **S** e **V**) a media e desviación estándar son dous valores, excepto na representación de matices (**H** e **HCS**), que conteñen catro valores medios (norma e ángulo) e dous valores de desviación estándar. O total de características (media e desviación estándar) é a suma: 10.

As características baseadas en estimacións de complexidade forman parte de tres grupos:

- Os métodos de compresión **JPEG** e **fractal**: combinación de 21 funcións.
- Os métodos das regras de **Zipf**: mínimo de 16 funcións compatibles.
- O recuento de caixas baseado na dimensión **fractal** e dos bordos inspirado en **Taylor et al.** (1999).

Considerouse na súa totalidade unha categoría composta por 53 métricas:

- 21 características de erro de compresión medidas usando os métodos **JPEG** e **fractal**

Autora: **Iria María Santos López**

Titor: **Juan Jesús Romero Cardalda**

Directores: **Adrián Carballal Mato** e **Juan Jesús Romero Cardalda**

Universidade da Coruña

- 16 funcións extraídas do enfoque baseado en **Zipf** para os rangos de frecuencia e a relación de escala
- 6 características da dimensión **fractal**
- 10 características baseadas na media e na desviación estándar

De cada imaxe deriváronse seis filtros adicionais (a incorporación horizontal e vertical dos filtros **Canny** e **Sobel**) por iso 371 métricas constitúen o grupo resultante de funcións: 53 funcións para 7 filtros.

4.3 Arquitectura proposta

As redes de neuronas artificiais están baseadas nunha analogía coa estrutura física do cerebro. Intentan empregar moitos procesadores sinxelos de forma cooperativa. A unión dos modelos pode formar distintas arquitecturas, pero deben ser adestrados para un suposto determinado por unha regra ou algoritmo determinado, que é o encargado de modificar os parámetros da rede para que se adapte aos obxectivos.

Existen dous tipos de adestramento: supervisado (un axente externo controla o proceso de aprendizaxe) e sen supervisión (a propia rede é a que desenvolve regras que organizan o conxunto de adestramento).

Tras finalizar o adestramento a rede debería poder xeneralizar extrapolando os datos aprendidos a novos padróns de entrada. Para iso o conxunto de adestramento debe ser amplo e representativo.

Unha das arquitecturas básicas máis empregadas é a **MLP** formulada por **Rosenblatt** (1958). A arquitectura ten un deseño de rede neuronal de varias capas composta por unha de entrada, unha de saída e varias intermedias (capas ocultas).

Dentro das redes neuronais existen dúas grandes tipoloxías: as **feedforward** (**FFNN**) e as **recorrentes** (**RNN**). Nas **FFNN** a información móvese nunha soa dirección e nas **RNN** forman un bucle na arquitectura. Ao se trataren de conexións circulares a optimización da rede **RNN** sería computacionalmente máis complexo, por iso buscamos optimizar a topoloxía dunha **MLP** integrada na **FFNN**.

Actualmente a tendencia principal é cara á aprendizaxe profunda onde o número de capas ocultas pode variar de un a centos. A capacidade de manipular grandes cantidades de información reduciu a barreira do tempo para adestrar estas tipoloxías.

Autora: **Iria María Santos López**

Titor: **Juan Jesús Romero Cardalda**

Directores: **Adrián Carballal Mato** e **Juan Jesús Romero Cardalda**

Universidade da Coruña

Neste caso preténdese o contrario do habitual, identificar a rede neuronal mínima completamente conectada que ofrezca resultados similares a outras máis grandes sen ter exceso de adestramento ou perda na capacidade de xeneralización.

Blum (1992) estableceu a topoloxía ideal nun máximo de 186 unidades ocultas. A nosa topoloxía debería ter 371 entradas, unha por cada característica detallada por **Machado et al.** (2015) e unha neurona única na capa de saída (problema binario). Se temos en conta o número de conexións existentes no adestramento sería dun máximo de 800 conexións.

Dada a diferenza entre as propostas de **Blum** (1992) e **Machado et al.** (2015) tómanse ambos os valores para establecer o rango de busca da topoloxía óptima do valor 2 para 2 unidades ocultas necesarias.

Ademais da topoloxía básica tamén deben definirse os parámetros. Algúns son estáticos: función de saída (identidade), función de actualización, inicial de pesos e algoritmo de adestramento. Tamén hai outros que deben identificarse mediante un proceso de hiperparametrización. Todos os parámetros seleccionados para a hiperparametrización están relacionados con valores numéricos que se seleccionan por proba-erro, excepto a función de transferencia, que queda estática xunto coa función de identidade. Non se estudan outras alternativas para non aumentar o tempo de cálculo.

Para identificar a topoloxía óptima e os demais parámetros utilízanse 50 probas independentes con cada combinación para que os resultados sexan relevantes. Os resultados son a media dos obtidos nas 50 carreiras.

A validación cruzada é o método científico empregado, porque outorga ao modelo a oportunidade de adestrar en múltiples divisións de probas de tren, dando unha idea clara do que faría en resultados invisibles⁶.

Escolleuse repetir experimentos de *5-fold cross-validation*,⁷ que é un número suficiente de veces, e observar o comportamento dos modelos. Se non existe

6 Outros métodos de clasificación coñecidos son a análise discriminante, o algoritmo kNN, as árbores de clasificación, a árbore C5.0, a máquina de soporte vectorial (SVM), as redes de neuronas artificiais, o clasificador Bayesiano e o combinado de clasificadores [Parra, 2017 – dispoñible *online* en: <https://bookdown.org/content/2274/portada.html> (11/03/2021)]

7 A validación cruzada é esencial para determinar a calidade da predición do modelo. Existen varios tipos de validación cruzada: *k-fold CV*, *repeated k-fold CV*, *leave-one-out CV (LOOVC)* e o tipo *Bootstrap CV* (Feng et al., 2005; Kuhn e Johnson, 2013).

variabilidade significativa nos resultados non é necesario aplicar unha validación cruzada distinta.

O obxectivo é atopar a topoloxía que ofrezca os mellores resultados para os catro conxuntos de datos de entrada (5%, 10%, 15% e 20%). Determinouse que 16 unidades na capa oculta son suficientes cumprindo as afirmacións de **Blum** (1992); pero tendo 371 características de entrada a topoloxía estaría composta por 6.307 conexións. Na Figura 11 móstranse os resultados obtidos para cada un dos catro conxuntos porcentuais empregando esas 371 características.

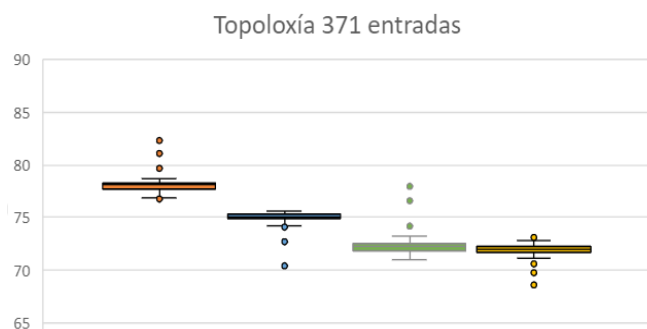


Figura 11: Trazo gráfico que mostra a distribución dos resultados das catro combinacións porcentuais en cada extremo cos datos da topoloxía con 371 entradas. As cores equivalen aos distintos adestramentos: laranja (5%), azul (10%), verde (15%) e amarelo (20%).

Para diminuír as conexións, reduciuse o número de elementos na capa de entrada. Determináronse as características máis representativas mediante a análise de compoñentes principais (**PCA**). Utilizouse esta técnica por ser un método lineal, xa que as outras opcións non son lineais e, polo tanto, non son tan boas opcións pola topoloxía de capa única proposta. Pódense ver os resultados dos distintos conxuntos de adestramento (5%, 10%, 15% e 20%) segundo o número de características seleccionados pola **PCA** na Táboa 2.

Táboa 2: Resultados obtidos adestrando a topoloxía óptima empregando características de entrada reducidas seleccionadas pola PCA. Os valores mostrados corresponde á media de 50 carreiras independentes. As cores equivalen aos distintos adestramentos: laranxa (5%), azul (10%), verde (15%) e amarelo (20%).

Características	Resultado (%)	Características	Resultado (%)	Características	Resultado (%)	Características	Resultado (%)
5	75'45	6	77'43	7	77'04	8	78'22
	71'89		72'38		72'57		73'81
	67'92		69'57		69'88		70'07
	67'26		68'74		69'17		70'04
9	77'45	10	79'38	11	78'85	12	79'95
	73'34		74'98		74'66		74'65
	69'64		70'01		70'73		71'38
	70'1		70'79		70'78		70'85
13	79'86	14	79'89	15	79'70	16	80'27
	75'02		75'58		75'35		75'4
	72'24		72'16		72'76		73
	70'84		70'79		71'34		71'26
17	80'09	18	80'29	19	79'95	20	79'33
	75'62		75'61		75'65		75'78
	73'06		74'84		74'29		74'88
	71'16		71'19		72'49		72'65

Tras observar os resultados chegouse á conclusión de que nun punto intermedio entre 10% e 15% hai un conxunto de imaxes en cada extremo que representa o desafío cognitivo para a rede neuronal. No conxunto da columna do 5% existe unha gran diferenza co resto, polo que é lóxico que a rede poda facer máis precisa a dicotomía.

Autora: **Iria María Santos López**

Titor: **Juan Jesús Romero Cardalda**

Directores: **Adrián Carballal Mato** e **Juan Jesús Romero Cardalda**

Universidade da Coruña

Realízase un estudo para determinar se as mostras independentes seguen ou non unha distribución normal mediante unha proba de **Kolmogorov–Smirnov**. Determínase que se necesita empregar unha proba non paramétrica de **Mann–Whitney U**⁸, recomendábel nos casos en que a comparación sexa só entre dúas mostras.

Case todos os enfoques detallados empregan modelos baseados en *SVM* e por iso foi comprobado se o subconxunto de 18 características obtén resultados similares con outros modelos de adestramento. Neste caso faise tamén un proceso de hiperparametrización para atopar a mellor configuración posíbel. Nos grupos do 10% ao 15% o salto é máis agudo, polo que o problema da distinción dos conxuntos é independente do modelo empregado. O modelo baseado en *SVM* usa un modelo lineal, polo que cumpre con outros supostos mencionados.

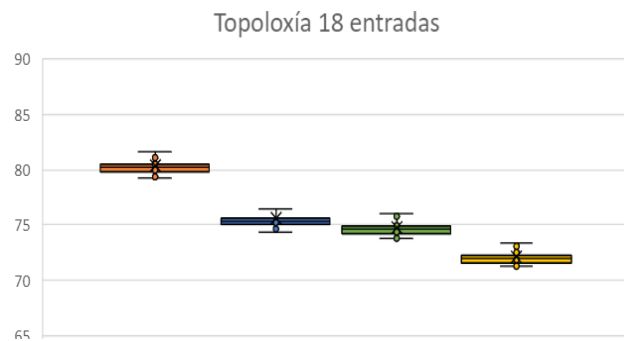


Figura 12: Trazo gráfico que mostra a distribución dos resultados das catro combinacións porcentuais en cada extremo coa topoloxía óptima de 18 entradas determinadas pola PCA. As cores equivalen aos distintos adestramentos: laranxa (5%), azul (10%), verde (15%) e amarelo (20%).

A topoloxía óptima de **MLP** contaría con 18 entradas, 16 unidades ocultas e unha saída; o que significaría un total de 340 conexións que representan unha relación de 9:1. Os resultados obtidos para esta topoloxía de 18 entradas nos diferentes conxuntos de porcentaxes poden verse na Figura 12. Porén, non cumpre co criterio de **Blum** (1992), xa que para 18 unidades de entrada e unha de saída deberíamos obter

⁸ A proba de Mann–Whitney U é unha proba estatística non paramétrica que comproba as diferenzas entre dous grupos nunha única variable ordinal sen distribución específica (Mann e Whitney, 1947). É diferente da proba de Kruskal–Wallis (que se utiliza no artigo 3), que avalía as diferenzas entre tres ou máis grupos de mostraxe independentes nunha única variable continua normalmente non distribuída, polo que se trata dunha extensión da anterior, pero máis xeneralizada (Kruskal e Wallis, 1952).

como máximo 9 unidades ocultas e non 16. O criterio de **Blum** (1992), deste xeito, é máis adecuado para redes **MLP** conectadas localmente.

Polo tanto, con esta investigación logrouse identificar unha rede mínima e compacta que utiliza só valores de entrada relacionados coa complexidade da imaxe onde a capacidade de xeneralización e rendemento se manteñen respecto a redes máis grandes. Sería interesante intentar replicar a metodoloxía con outras redes máis complexas como o modelo **RNN**.

Autora: **Iria María Santos López**

Titor: **Juan Jesús Romero Cardalda**

Directores: **Adrián Carballal Mato e Juan Jesús Romero Cardalda**

Universidade da Coruña

5 Comparison of Outlier-Tolerant Models for Measuring Visual Complexity

Como xa falamos con anterioridade, a cuantificación da complexidade visual dunha imaxe é unha tarefa que resulta moi útil, á par que complexa. Moitos factores, caso da exposición previa a certos estímulos visuais dados, afectan á percepción humana da complexidade visual. Autores como **Frijda** (1989) ou **Blood** e **Zatorre** (2001) tamén afirman que a experiencia estética é afectiva ou incluso emocional. **Leder et al.** (2004) crearon un modelo que diferencia entre emoción estética e xuízos estéticos como dous tipos de saída, ofrecendo aos investigadores a flexibilidade para consideraren e controlaren variables como o nivel de coñecemento e o estado afectivo dos destinatarios. **Graf** e **Landwehr** (2015) tamén contribuíron cun modelo que permite formar preferencias estéticas. Pero foi **Berlyne** (1970) quen proporcionou unha explicación psicobiolóxica adecuada dos efectos da complexidade na preferencia, tras analizar o resultado da interacción do sistema de recompensa e aversión. Con isto, comprendeu que os ditos sistemas levarían ás persoas a preferiren niveis intermedios de complexidade, definidos por características como a regularidade do padrón, a cantidade de elementos, a heteroxeneidade ou a irregularidade de formas. Isto débese a que as persoas buscan manter o nivel de excitación preferido e valores inferiores ou superiores fan que se perda o interese.

Existen dúas características, que inflúen na percepción da complexidade, consideradas dende vai moito tempo como determinantes da beleza: a orde e a variedade. Xa se viu que **Birkhoff** (2013) definiu unha función destes tres elementos (fórmula 1).

Onde M é o valor estético, O a orde e C a complexidade. Podemos obter desta ecuación a certeza de que a menor complexidade, maior é a beleza. Pola contra, tamén vimos que nos estudos de **Eysenck** (1942) a correlación cun grupo de humanos suxeriron que tanto a orde como a complexidade contribúen de forma positiva á beleza. **Van Geert** e **Wagemans** (2020) tamén apoian a importancia da complexidade e da orde na predición da apreciación estética.

Autora: **Iria María Santos López**

Titor: **Juan Jesús Romero Cardalda**

Directores: **Adrián Carballal Mato** e **Juan Jesús Romero Cardalda**

Universidade da Coruña

5.1 *Dataset* e traballos previos

Os primeiros métodos para predicir a complexidade visual baseáronse no número de elementos e na súa regularidade e heteroxeneidade. Formáronse por diferentes conxuntos de polígonos, que permitían contar de forma manual os elementos contidos. **Lempel e Ziv** (1976) foron os que formularon un algoritmo novo para medir a complexidade visual, que destaca porque o crearon a partir do programa de ordenador máis pequeno necesario para almacenar e producir unha imaxe. Este método segue a teoría de que a lonxitude mínima de código necesaria para describir unha imaxe é unha medida adecuada de complexidade. Outros estudos sobre estética experimental publicados recentemente centráronse en investigar como o tempo de exposición do estímulo afecta ao procesamento de imaxes. O método máis empregado é a avaliación de imaxes por un grupo de humanos. **Forsythe et al.** (2011) analizaron a correlación entre o valor de complexidade dado polas persoas a 800 estímulos visuais e os resultados das medidas de compresión **JPEG** e **GIF**, e unha medida de detección perimetral. Os resultados mostraron que as tres medidas estaban correlacionadas coa complexidade xulgada pola xente. A compresión **GIF** mostrou a correlación máis forte ($R\text{-Squared} = 0'5476$). Tamén no campo da psicoloxía estética se desenvolveron traballos nos que distintas imaxes foron avaliadas por humanos segundo criterios exclusivamente estéticos.

Marin e Leder (2013) estudaron diferentes medidas computacionais relacionadas coa clasificación de compacidade de diferentes materiais e descubriron que a compresión **TIFF** ($R\text{-Squared} = 0'2809$) e a **JPEG** ($R\text{-Squared} = 0'2704$) tiñan mellor correlación que a detección perimetral con valores de complexidade subxectiva. **Nadal e Chatterjee** (2019) descubriron que as persoas difiren significativamente na medida na que as características visuais inflúen no seu gusto.

Forsythe et al. (2011) basearon o seu traballo no conxunto de datos creado por **Cela-Conde et al.** (2009) para medir a correlación entre a complexidade visual percibida polo ser humano e as diferentes medias obtidas por computador.

Posteriormente, **Machado et al.** (2015) propuxeron as novas estimacións de complexidade baseadas no erro de compresión e na lei de **Zipf**. O obxectivo de **Machado et al.** (2015) foi mellorar os resultados na predición do valor de complexidade percibido polos humanos mediante un método computacional. Tamén intentaron avaliar os seus resultados de correlación de percepción individual

Autora: **Iria María Santos López**

Titor: **Juan Jesús Romero Cardalda**

Directores: **Adrián Carballal Mato** e **Juan Jesús Romero Cardalda**

Universidade da Coruña

combinando as súas características vectoriais con redes de neuronas artificiais cun resultado de *R-Squared* de ata o 0'6939. Empregaron redes de neuronas artificiais e métricas xeradas por ordenador con *machine learning* para predicir a complexidade visual das imaxes. **Carballal et al.** (2018) estudaron a complexidade visual usando as características da imaxe e unha rede de neuronas artificiais para clasificar un conxunto de imaxes e discriminar automaticamente entre fotografías e pinturas.

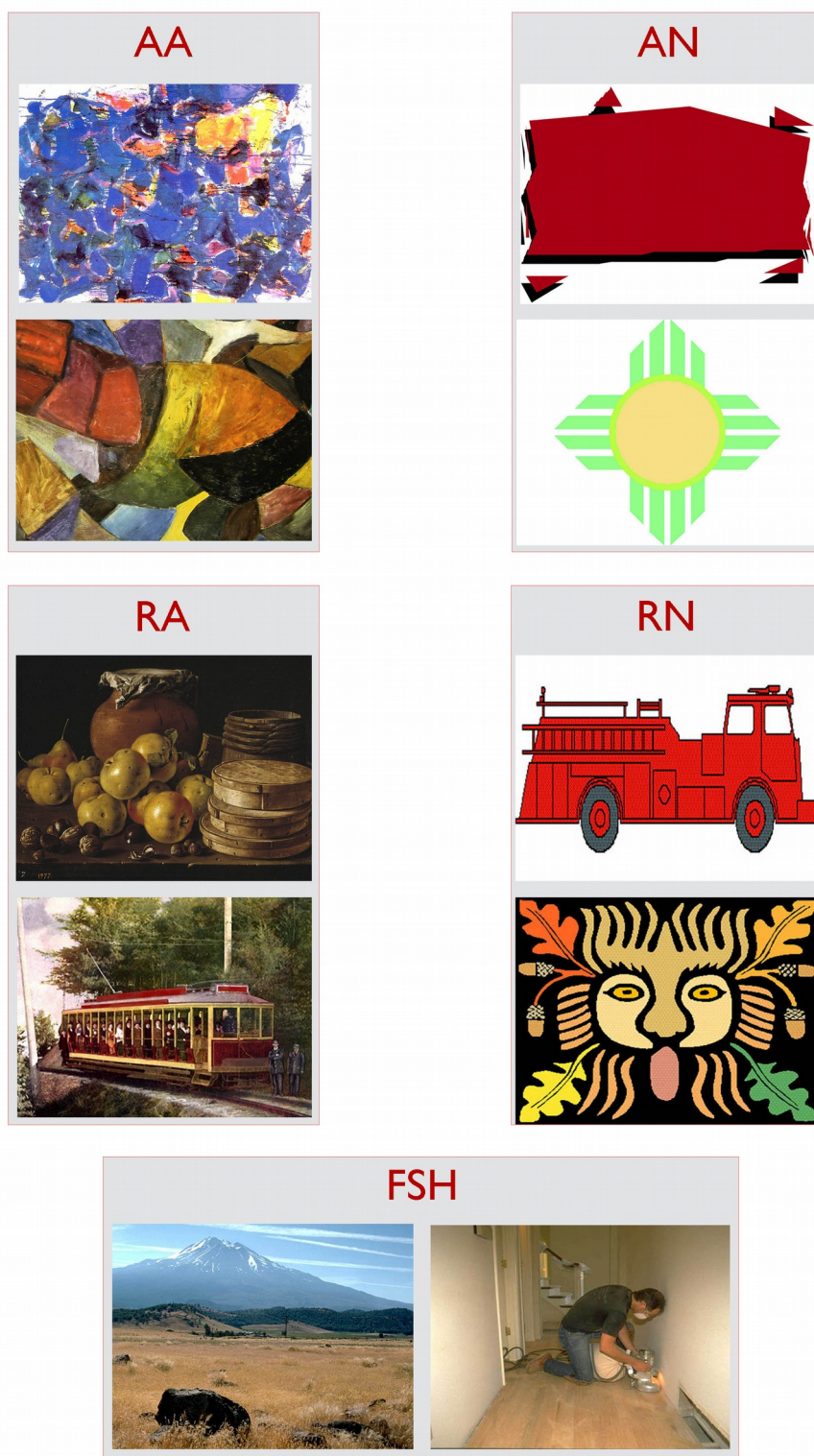


Figura 13: Dous exemplos das imaxes que se inclúen nas cinco categorías presentes no dataset de **Forsythe et al.** (2011) empregado neste experimento: abstracto artístico (AA), abstracto non artístico (AN), representativo artístico (RA), representativo non artístico (RN) e fotografías creadas por seres humanos (FSH).

Autora: Iria María Santos López
Titor: Juan Jesús Romero Cardalda
Directores: **Adrián Carballal Mato** e **Juan Jesús Romero Cardalda**
Universidade da Coruña

Neste traballo, tómase como datos de referencia os empregados por **Forsythe et al.** (2011), **Machado et al.** (2015) e **Fernandez-Lozano et al.** (2019). Probáronse varios modelos computacionais diferentes seguindo un deseño experimental robusto para a análise de regresión avaliando o rendemento en diferentes experimentos de modelos computacionais de aprendizaxe automática múltiple e comparando as diferenzas estatísticas dos resultados. O rendemento dos modelos foi avaliado co uso de R^2 a través de 50 probas dun experimento de validación cruzada de 10 veces en conxuntos distintos. A validación cruzada ten o propósito de probar a capacidade do modelo para predicir novos datos que non se utilizan no adestramento para que se poidan detectar problemas como o exceso de tráfico ou o nesgo de selección, e proporciona unha visión de como o modelo se xeneraliza nun conxunto de datos separados. Os resultados foron avaliados analizando o conxunto completo de variables e elimináronse as atípicas para validar o pre-procesamento.

O conxunto de datos empregado foi o do estudo de **Forsythe et al.** (2011) que proviña de **Cela-Conde et al.** (2009) e contiña preto de 1.500 imaxes dixitalizadas. Deste conxunto completo o total final (**Forsythe et al.**, 2011) foron 800 imaxes divididas en cinco categorías: representativo artístico (RA), abstracto artístico (AA), representativo non artístico (RN), abstracto non artístico (AN) e fotografías creadas polo ser humano (FSH). Como pode verse na Figura 13 as imaxes incluídas nas AA son pinturas abstractas; as AN son figuras vectoriais abstractas; as RA son pinturas, que mostran unha representación clara da realidade; as RN son figuras vectoriais, que representan a realidade; e as FSH son fotografías.

No traballo de **Forsythe et al.** (2011) participaron 240 persoas onde un 46'6% eran homes e un 53'33% eran mulleres, que formaban parte da Universidade das Illas Baleares sen formación artística formal. As 800 imaxes foron divididas en conxuntos de 100 imaxes e cada conxunto se lles presentou de forma aleatoria aos grupos compostos por 30 persoas. As imaxes mostráronse aos participantes durante cinco segundos nunha pantalla de 400x225 cm (16:9) a través dun *PowerPoint* cos participantes a distancias entre 200 e 700 cm. Os resultados non teñen carácter público, pero foron proporcionados polos autores orixinais que colaboraron deste modo no avance da investigación.

Cada estímulo fixouse en 150 ppi cun tamaño de 9x12 centímetros. A escala de cor variouse maximizando as variables psicofísicas e a luminancia da estimación de 270 a 390 luxes. En certos casos, para unha anonimización óptima, a sinatura do autor

Autora: **Iria María Santos López**

Titor: **Juan Jesús Romero Cardalda**

Directores: **Adrián Carballal Mato** e **Juan Jesús Romero Cardalda**

Universidade da Coruña

eliminouse de forma uniforme. A cada participante déronlle uns estímulos pouco complicados ao comezo do experimento (unha icona) e outro de gran complexidade (unha cidade) como exemplos xunto cunha descrición de detalles ou dificultade (**Snodgrass** e **Vanderwart**, 1980). Os participantes tiveron que clasificar individualmente as imaxes nunha escala de Likert (de 1 a 5), onde 5 eran imaxes complexas e 1 moi sinxelas. O resultado para 5 imaxes da categoría fotografías e unha da categoría AN os valores foron atípicos dentro dalgún dos seguintes criterios: distancias de Cook, residuais estudados e altos puntos de apalancamento; por iso foron suprimidas.

5.2 Modelos estudados

Realizáronse varios experimentos para seleccionar o mellor modelo usando **R** e o paquete **RRegrs**, **MATLAB R** e **CGS** baseado en **Java**. Os métodos empregados foron os seguintes: *Elastic Net (ENET)*, selección de funcións aprendizaxe múltiple do núcleo (*FSMKL*), regresión parcial de mínimos cadrados (*PLS*), *Random Forest (RF)*, eliminación de características recursivas do bosque aleatorio (*RF-RFE*), soporte de máquinas vectoriais con funcións radiais (*SVM-Radial*) e un algoritmo xenético híbrido baseado en modelo de regresión chamado *Correlation by Genetic Search (CGS)*.

- Regresión *ENET*: combina a potencia dunha regresión de **Ridge** (elimina as variables non relevantes) e **Lasso** (reduce o número ao máis próximo a 0) nun único algoritmo mediante un hiperparámetro que controla as variables analizadas. Trátase dun tipo especial de algoritmo de regresión pertencente ao grupo de algoritmos de regularización. Este modelo foi creado como crítica ao modelo **Lasso**, que tiña unha capacidade para seleccionar as variables moi dependente dos datos analizados e alta variabilidade en función dos exemplos empregados no conxunto de datos.
- *FSMKL*: modelo baseado en núcleos moi empregado para resolver diversos problemas de diferentes campos. Adoita funcionar en ambientes de alta dimensionalidade e ruidosos. Os diferentes subgrupos de variables extraéronse das imaxes individuais codificándose en núcleos e combináronse os núcleos para resolver un problema de regresión co uso de *SVM*. Realizouse un proceso interno de selección de funcións en cada subconxunto escollendo as variables e subconxuntos máis relevantes.

Autora: **Iria María Santos López**

Titor: **Juan Jesús Romero Cardalda**

Directores: **Adrián Carballal Mato** e **Juan Jesús Romero Cardalda**

Universidade da Coruña

- *PLS*: modelo que consiste en reducir ao máximo o número de variables, seleccionar as de menor correlación entre elas e realizar unha regresión de mínimos cadrados só no subconxunto reducido. Emprégase moito cando se analizan datos de alta dimensionalidade e correlacionados entre si. O modelo usa combinacións lineares das variables calculadas cunha análise de compoñentes principais (PCA) seleccionando as proxeccións que mellor describen a covarianza dos datos e a variable dependente.
- *RF*: técnica capaz de resolver problemas de regresión e clasificación coa combinación de árbores de decisión sinxelas sen correlación con distintos parámetros para cada árbore e tamén diferentes conxuntos de datos. O resultado obtido é a media dos resultados de todas as árbores de decisión.
- *RF-RFE*: é unha combinación da *RF* anteriormente mencionada e unha fase de selección de funcións externa coñecida como eliminación de funcións recursivas (*RFE*). A *RFE* extrae en cada iteración do algoritmo de *RF* a importancia das variables na predición final e elimina as que tiveron menos peso no proceso para atopar o subconxunto de variables de mellor resultado.
- *SVM-Radial*: modelo de máquinas vectoriais de soporte (*SVM*) para problemas de regresión cun núcleo de función de base radial (*RBF*), que tenta identificar o hiperplano que maximiza a marxe de separación. Para os problemas de regresión débese engadir un factor de tolerancia de erros, posto que as saídas son números reais. *SVM*, pola súa banda, basea o funcionamento na construción dun hiperplano óptimo en forma de superficie de decisión, no cal a marxe de erro é mínima na estimación da mellor regresión. A potencia do *SVM* atópase en que permite aumentar a dimensionalidade dos datos e transformar o espazo de entrada nun máis grande sen aumento do custo computacional.
- *CGS*: algoritmo xenético híbrido baseado en modelos de regresión. Foi empregado por **Pazos-Perez et al.** (2018) para predicir o crecemento urbano vertical en Tokio baseándose en datos macroeconómicos. Tamén o empregaron **Carballal et al.** (2019) para predicir a estética. O algoritmo xenético emprega operadores matemáticos para refactorizar as variables de entrada e atopar unha solución adecuada.

Autora: **Iria María Santos López**

Titor: **Juan Jesús Romero Cardalda**

Directores: **Adrián Carballal Mato** e **Juan Jesús Romero Cardalda**

Universidade da Coruña

Todos os modelos tiveron resultados mellores ou máis estables ao analizar o conxunto de datos do que se eliminaron os valores atípicos excepto con *ENET*.

A mellora no modelo *FSMKL* foi significativa, ao poder aumentar o seu rendemento segundo R^2 en varios puntos porcentuais. *CGS*, o segundo modelo gañador, foi o que presentou un comportamento con maior variabilidade estatística ao longo das 50 repeticións, pero conseguiu superar a todos os demais dun xeito estatisticamente significativo; ofrecendo unha maior variabilidade nos dous casos (con valores atípicos e sen eles).

Os catro mellores modelos (*RF*, *RF-RFE*, *SVM-Radial* e *CGS*) separáronse nos dous problemas (conxunto completo e conxunto prescindindo dos valores atípicos) e empregáronse para unha proba emparellada de **Kruskal-Wallis** cun nivel significativo de $\alpha = 0'05$ e asumindo a hipótese nula de que os resultados fosen estatisticamente iguais.

A diferenza do método *FSMKL*, onde o uso do filtro de detección de bordos **Canny** e a canle de cor de saturación parecía ter máis relevancia, co método *CGS* tanto esa canle de cor en particular como o uso ou non de filtros de detección de bordos non eran tanto unha prioridade. *FSMKL* puido usar as pequenas diferenzas entre as métricas (que estaban moi correlacionadas) para obter correlacións da orde de 0'71. No caso de *CGS*, o método parecía atopar mellores aproximacións usando grupos de funcións cun menor nivel de correlación interna en comparación con *FSMKL*.

O resultado obtido foi que o modelo *CGS* (método híbrido para a creación de modelos de regresión múltiple baseados na maximización da correlación) acadou a mellor puntuación de rendemento entre as 50 diferentes corridas realizadas para cada un dos conxuntos de datos empregados.

O método *CGS* tivo os mellores resultados de correlación en termos de *R-Squared*, que foron corroborados estatisticamente a través da proba parella de **Kruskal-Wallis**. Este método tamén se estudou eliminando do conxunto de estudos imaxes xa identificadas coma valores atípicos en estudos anteriores, así coma o resto dos métodos propostos. Os resultados mostran que o método permanece estable en termos estatísticos e as métricas que o método precisa para obter estes resultados apenas varían. Do mesmo xeito que os valores *R-Squared*, onde a variabilidade radial

Autora: **Iria María Santos López**

Titor: **Juan Jesús Romero Cardalda**

Directores: **Adrián Carballal Mato** e **Juan Jesús Romero Cardalda**

Universidade da Coruña

SVM era mínima, *CGS* tamén foi estatisticamente mellor en termos de *RMSE*, aínda que este é un método con maior variabilidade.

Estas evidencias mostran a capacidade do modelo para identificar métricas que non dependen de posibles valores atípicos no conxunto de entrada. A elección do método *CGS*, non só en función dos resultados obtidos, é adecuada xa que os datos de entrada teñen unha dimensión relativamente pequena (329 características), o que permite usar poboacións igualmente pequenas.

6 Conclusións

Nesta tese de doutoramento por compendio preséntanse tres traballos nos que se empregan técnicas adaptativas co fin de predicir a preferencia visual, ben dunha forma directa, ben mediante outras variábeis subxectivas como é a percepción de complexidade visual.

Se ben, o primeiro dos artigos trátase dun estado da arte que non engloba todos os contidos relacionados coa **IA** en artes visuais, senón só os relativos a redes de neuronas artificiais e *deep learning*, isto débese a que xa existen investigacións previas doutros autores que tratan o tema de forma máis xeral. No noso caso, ao longo da investigación seleccionouse todo este contido, pero posteriormente tan só se tiveron en conta os artigos relacionados con redes de neuronas artificiais e *deep learning*, posto que non existían revisións previas ao respecto tan específicas, o cal consideramos de interese. Queda constancia do interese científico na posibilidade de publicación do artigo nunha revista cun índice de impacto que abrangue o cuartil **Q1**.

Desta revisión podemos concluír dous datos importantes á hora de levar a cabo os experimentos posteriores: existen poucos artigos enfocados á complexidade visual (tan só un dentro deste estado da arte) e, pola contra, houbo un auxe nos últimos anos nos avances relacionados coa avaliación da estética.

Considerando estes dous datos:

- Se outros autores están tratando o tema da avaliación e, por tanto, da predición de estética visual nas obras de arte, non cabe dúbida de que se trata dun campo de relevancia. Neste sentido, buscar a topoloxía mínima capaz de realizar estas tarefas é un factor chave para a mellora dos procesos.
- A falta de traballos de predición de complexidade visual pode mostrar unha baixa relevancia do tema. No entanto, como pode verse en **Johnson et al.** (2019), a complexidade visual é o factor máis nomeado nos estudos sobre a psicoloxía da estética en relación coa predición do valor estético. É por iso que o que consideramos moi relevante á par que complexo dado o seu alto compoñente subxectivo.

Volvendo ás preguntas presentadas na introdución (ten sentido crer que un enfoque computacional da estética nos axudará a entendela?; merece a pena investigar ferramentas e contornas computacionais que poidan contribuír a temas

Autora: **Iria María Santos López**

Titor: **Juan Jesús Romero Cardalda**

Directores: **Adrián Carballal Mato** e **Juan Jesús Romero Cardalda**

Universidade da Coruña

relacionados coa estética?; é factible crear programas que aporten contidos estéticos de relevancia?), coa lectura desta tese, a resposta clara a todas as preguntas é un si rotundo, xa que, como vimos, pode servir para innumerables aplicacións, dende psicolóxicas a comerciais.

A hipótese presentada sobre a posibilidade de predicir a preferencia de usuario mediante técnicas de *soft computing* é algo que ten verdadeira importancia. Cos datos mostrados tan só no estado da arte (Artigo 1), debeu quedar claro que esta posibilidade é real. Por iso, fomos un paso máis alá, realizando prácticas para poder obter esta predición estética da mellor forma posible con menos custos computacionais, económicos, temporais, e con resultados satisfactorios e de peso.

Logrouse, finalmente, identificar unha rede mínima compacta a través do uso de valores de entrada relacionados coa xa famosa complexidade da imaxe e con capacidade de xeneralización e rendemento (Artigo 2). Descubriuse tamén que o método híbrido para a creación de modelos de regresión múltiple baseados na maximización da correlación (CGS) é o que obtén mellor rendemento na predición de estética visual (Artigo 3). Por iso, considérase idóneo o emprego de ambas as técnicas de *machine learning* para esta problemática.

Queda pendente para investigacións futuras replicar a metodoloxía da *MLP* en redes máis complexas como o modelo **RNN**. Tamén resulta de interese para investigacións futuras, sobre todo en posibles aplicacións para *márketing* dixital e *e-commerce*, o cálculo de predicións directas sobre o gusto estético de usuarios individuais, xa que os estudos levados a cabo ata o momento foron sobre o gusto estético de conxuntos de persoas.

Autora: **Iria María Santos López**

Titor: **Juan Jesús Romero Cardalda**

Directores: **Adrián Carballal Mato e Juan Jesús Romero Cardalda**

Universidade da Coruña

Referencias

1. Ahmed, M. A., Hasan, R. A., Ali, A. H., & Mohammed, M. A. (2019). The classification of the modern arabic poetry using machine learning. *Telkomnika*, 17(5), 2667-2674.
2. Alcolea, S. (1975). *Artes decorativas en la España cristiana:(siglos XI-XIX)*. Ed. Plus-Ultra.
3. Barron, F. (1963). Creativity and psychological health.
4. Batteaux, C. (1747). *Les Beaux-Arts réduits à un même principe*. chez Durand.
5. Bazin, A. (2004). *The ontology of the photographic image* (pp. 9-16). University of California Press.
6. Ben-Bassat, T., Meyer, J., & Tractinsky, N. (2006). Economic and subjective measures of the perceived value of aesthetics and usability. *ACM Transactions on Computer-Human Interaction (TOCHI)*, 13(2), 210-234.
7. Berlyne, D. E. (1970). Novelty, complexity, and hedonic value. *Perception & psychophysics*, 8(5), 279-286.
8. Bhandari, U., Chang, K., & Neben, T. (2019). Understanding the impact of perceived visual aesthetics on user evaluations: An emotional perspective. *Information & Management*, 56(1), 85-93.
9. Bianco, S., Mazzini, D., & Schettini, R. (2017, September). Deep multibranch neural network for painting categorization. In *International conference on image analysis and processing* (pp. 414-423). Springer, Cham.
10. Birkhoff, G. D. (2013). *Aesthetic measure*. Harvard University Press.
11. Blair, A. (2019, April). Adversarial evolution and deep learning—how does an artist play with our visual system?. In *International Conference on Computational Intelligence in Music, Sound, Art and Design (Part of EvoStar)* (pp. 18-34). Springer, Cham.
12. Blood, A. J., & Zatorre, R. J. (2001). Intensely pleasurable responses to music correlate with activity in brain regions implicated in reward and emotion. *Proceedings of the national academy of sciences*, 98(20), 11818-11823.
13. Blum, A. (1992). *Neural networks in C++ an object-oriented framework for building connectionist systems*. John Wiley & Sons, Inc..
14. Bory, P. (2019). Deep new: The shifting narratives of artificial intelligence from Deep Blue to AlphaGo. *Convergence*, 25(4), 627-642.
15. Bradford, J. H., & Côté-Laurence, P. (1995). An application of artificial intelligence to the choreography of dance. *Computers and the Humanities*, 29(4), 233-240.
16. Bringsjord, S., & Ferrucci, D. (1999). *Artificial intelligence and literary creativity: Inside the mind of brutus, a storytelling machine*. Psychology Press.

Autora: **Iria María Santos López**

Titor: **Juan Jesús Romero Cardalda**

Directores: **Adrián Carballal Mato** e **Juan Jesús Romero Cardalda**

Universidade da Coruña

17. Briot, J. P., Hadjeres, G., & Pachet, F. D. (2017). Deep learning techniques for music generation--a survey. *arXiv preprint arXiv:1709.01620*.
18. Cage, J. (1991). An autobiographical statement. *Southwest Review*, 76(1), 59-76.
19. Canudo, R. (1911). Manifiesto de las siete artes. *Textos y Manifiestos del Cine. Estética. Escuelas. Movimientos. Disciplinas. Innovaciones*.
20. Carballal, A., Fernandez-Lozano, C., Heras, J., & Romero, J. (2019). Transfer learning features for predicting aesthetics through a novel hybrid machine learning method. *Neural Computing and Applications*, 1-12.
21. Carballal, A., Fernandez-Lozano, C., Rodriguez-Fernandez, N., Santos, I., & Romero, J. (2020). Comparison of outlier-tolerant models for measuring visual complexity. *Entropy*, 22(4), 488.
22. Carballal, A., Perez, R., Santos, A., & Castro, L. (2014, April). A complexity approach for identifying aesthetic composite landscapes. In *International Conference on Evolutionary and Biologically Inspired Music and Art* (pp. 50-61). Springer, Berlin, Heidelberg.
23. Carballal, A., Santos, A., Romero, J., Machado, P., Correia, J., & Castro, L. (2018). Distinguishing paintings from photographs by complexity estimates. *Neural Computing and Applications*, 30(6), 1957-1969.
24. Cardoso, A., & Romero, J. J. (2007). Creatividad Computacional. *Inteligencia Artificial y Computación Avanzada*, 149 – 176.
25. Castellano, G., & Vessio, G. (2020). Towards a tool for visual link retrieval and knowledge discovery in painting datasets. *Italian Research Conference on Digital Libraries* (pp. 105-110). Springer.
26. Cela-Conde, C. J., Ayala, F. J., Munar, E., Maestú, F., Nadal, M., Capó, M. A., ... & Marty, G. (2009). Sex-related similarities and differences in the neural correlates of beauty. *Proceedings of the National Academy of Sciences*, 106(10), 3847-3852.
27. Christiansen, H. C. (Ed.). (2000). *Comics & culture: analytical and theoretical approaches to comics*. Museum Tusulanum Press.
28. Chok, N. S. (2010). *Pearson's versus Spearman's and Kendall's correlation coefficients for continuous data* (Doctoral dissertation, University of Pittsburgh).
29. Correa, A. B. (Ed.). (1982). *Historia de las artes aplicadas e industriales en España*. Cátedra.
30. Cuenca, M., Lazcano, I., & Landabidea, X. (2010). Sobre ocio creativo: situación actual de las Ferias de Artes Escénicas.
31. Dannenberg, R. B. (2000). Artificial intelligence, machine learning, and music understanding. In *Proceedings of the 2000 Brazilian Symposium on Computer Music: Arquivos do Simpósio Brasileiro de Computação Musical (SBCM)*.

Autora: Iria María Santos López

Titor: Juan Jesús Romero Cardalda

Directores: Adrián Carballal Mato e Juan Jesús Romero Cardalda

Universidade da Coruña

32. Datta, R., Joshi, D., Li, J., & Wang, J. Z. (2006, May). Studying aesthetics in photographic images using a computational approach. In *European conference on computer vision* (pp. 288-301). Springer, Berlin, Heidelberg.
33. Datta, R., Li, J., & Wang, J. Z. (2008, October). Algorithmic inferencing of aesthetics and emotion in natural images: An exposition. In *2008 15th IEEE international conference on image processing* (pp. 105-108). IEEE.
34. De Mantaras, R. L., & Arcos, J. L. (2002). AI and music: From composition to expressive performance. *AI magazine*, 23(3), 43-43.
35. den Heijer, E., & Eiben, A. E. (2010, July). Using aesthetic measures to evolve art. In *IEEE Congress on Evolutionary Computation* (pp. 1-8). IEEE.
36. Desnoyer, M., & Wettergreen, D. (2010, August). Aesthetic image classification for autonomous agents. In *2010 20th International Conference on Pattern Recognition* (pp. 3452-3455). IEEE.
37. Dokic, J. (1998). Music, noise, silence: Some reflections on John cage. *Angelaki: Journal of the Theoretical Humanities*, 3(2), 103-112.
38. Dubray, D., & Laubrock, J. (2019, September). Deep CNN-based speech balloon detection and segmentation for comic books. In *2019 International Conference on Document Analysis and Recognition (ICDAR)* (pp. 1237-1243). IEEE.
39. Elgammal, A. (2019). AI is blurring the definition of artist: Advanced algorithms are using machine learning to create art autonomously. *American Scientist*, 107(1), 18-22.
40. Elgammal, A., Kang, Y., & Den Leeuw, M. (2018, April). Picasso, matisse, or a fake? Automated analysis of drawings at the stroke level for attribution and authentication. In *Proceedings of the AAAI Conference on Artificial Intelligence* (Vol. 32, No. 1).
41. Eysenck, H. J. (1942). The experimental study of the 'good Gestalt'—a new approach. *Psychological Review*, 49(4), 344.
42. Feduchi, L. (1946). *Historia del mueble*. Madrid: Aguado.
43. Feng, C. X. J., Yu, Z. G. S., Kingi, U., & Baig, M. P. (2005). Threefold vs. fivefold cross validation in one-hidden-layer and two-hidden-layer predictive neural network modeling of machining surface roughness data. *Journal of manufacturing systems*, 24(2), 93-107.
44. Feng, S. Y., & Ting, C. K. (2014, November). Painting using genetic algorithm with aesthetic evaluation of visual quality. In *International Conference on Technologies and Applications of Artificial Intelligence* (pp. 124-135). Springer, Cham.
45. Fernandez-Lozano, C., Carballal, A., Machado, P., Santos, A., & Romero, J. (2019). Visual complexity modelling based on image features fusion of multiple kernels. *PeerJ*, 7, e7075.
46. Fleming, J., & Honour, H. (1987). *Diccionario de las artes decorativas John Fleming y Hugh Honour; versión española y adaptación de María Luisa Balseiro*. Alianza Editorial.

Autora: Iria María Santos López

Titor: Juan Jesús Romero Cardalda

Directores: Adrián Carballal Mato e Juan Jesús Romero Cardalda

Universidade da Coruña

47. Forsythe, A., Nadal, M., Sheehy, N., Cela-Conde, C. J., & Sawey, M. (2011). Predicting beauty: fractal dimension and visual complexity in art. *British journal of psychology*, 102(1), 49-70.
48. Fray, S. B., Ortiz, P. A., & Naranjo, E. R. (2020). Formación académica en la industria gráfica de Riobamba-Ecuador. *Actas de Diseño*, (31).
49. Frijda, N. H. (1989). Aesthetic emotions and reality.
50. Galanter, P. (2012). Computational aesthetic evaluation: steps towards machine creativity. In *ACM SIGGRAPH 2012 Courses* (pp. 1-162).
51. Gao, F., Wang, Y., Li, P., Tan, M., Yu, J., & Zhu, Y. (2017). Deepsim: Deep similarity for image quality assessment. *Neurocomputing*, 257, 104-114.
52. Gero, J. S. (2014). *Artificial intelligence in design'91*. Butterworth-Heinemann.
53. Gero, J. S., & Maher, M. L. (2013). *Modeling creativity and knowledge-based creative design*. Psychology Press.
54. Gero, J. S., & Sudweeks, F. (Eds.). (2012). *Artificial Intelligence in Design'96*. Springer Science & Business Media.
55. Gips, J. (1975). *Shape grammars and their uses: artificial perception, shape generation and computer aesthetics*. Basel: Birkhäuser.
56. Gonthier, N., Ladjal, S., & Gousseau, Y. (2020). Multiple instance learning on deep features for weakly supervised object detection with extreme domain shifts. *arXiv preprint arXiv:2008.01178*.
57. Graf, L., & Landwehr, J. (2015). A dual-process perspective on fluency-based aesthetics: The pleasure-interest model of aesthetic liking. *Person. Soc. Psychol. Rev.*, (pp. 19, 395-410)
58. Greimas, A. J., Collins, F., & Perron, P. (1989). Figurative semiotics and the semiotics of the plastic arts. *New literary history*, 20(3), 627-649.
59. Gruber, A. (2000a). *Las Artes Decorativas en Europa. Primera Parte. Del Renacimiento al Barroco*. Madrid: Summa Artis. Vol. XLVI-I.
60. Gruber, A. (2000b). *Las Artes Decorativas en Europa. Segunda Parte. Del Neorrealismo al Art Déco*. Madrid: Summa Artis. Vol. XLVI-II.
61. Hartmann, J., Sutcliffe, A., & Angeli, A. (2008). Towards a theory of user judgment of aesthetics and user interface quality. *ACM Transactions on Computer-Human Interaction (TOCHI)*, 15(4), 1-30.
62. Hauke, J., Kossowski, T. (2011). *Comparison of values of Pearson's and Spearman's correlation coefficient on the same sets of data*.
63. Hicsonmez, S., Samet, N., Sener, F., & Duygulu, P. (2017, June). DRAW: Deep networks for Recognizing styles of Artists Who illustrate children's books. In *Proceedings of the 2017 ACM on international conference on multimedia retrieval* (pp. 338-346).

Autora: Iria María Santos López

Titor: Juan Jesús Romero Cardalda

Directores: Adrián Carballal Mato e Juan Jesús Romero Cardalda

Universidade da Coruña

64. Holland, J. (1996). Sistemas adaptativos complejos.
65. Hua, K. L., Ho, T. T., Jangtjik, K. A., Chen, Y. J., & Yeh, M. C. (2020). Artist-based painting classification using Markov random fields with convolution neural network. *Multimedia Tools and Applications*, 1-24.
66. Iges, J., Marie, J. L. & Fontán del Junco, M. (2016). *Escuchar con los ojos: Arte Sonoro en España, 1961-2016*. Fundación March.
67. Jafarpour, S., Polatkan, G., Brevdo, E., Hughes, S., Brasoveanu, A., & Daubechies, I. (2009, August). Stylistic analysis of paintings using wavelets and machine learning. In *2009 17th European Signal Processing Conference* (pp. 1220-1224). IEEE.
68. Jia, G., Li, P., & He, R. (2019). Theme aware aesthetic distribution prediction with full resolution photos. *arXiv preprint arXiv:1908.01308*.
69. Jin, X., Chi, J., Peng, S., Tian, Y., Ye, C., & Li, X. (2016, October). Deep image aesthetics classification using inception modules and fine-tuning connected layer. In *2016 8th International Conference on Wireless Communications & Signal Processing (WCSP)* (pp. 1-6). IEEE.
70. Johnson, C. G., McCormack, J., Santos, I., & Romero, J. (2019). Understanding aesthetics and fitness measures in evolutionary art systems. *Complexity*, 2019.
71. Joshi, B., Stewart, K., & Shapiro, D. (2017, July). Bringing impressionism to life with neural style transfer in come swim. In *Proceedings of the ACM SIGGRAPH digital production symposium* (pp. 1-5).
72. Karayev, S., Trentacoste, M., Han, H., Agarwala, A., Darrell, T., Hertzmann, A., & Winnemoeller, H. (2013). Recognizing image style. *arXiv preprint arXiv:1311.3715*.
73. Kasuya, E. (2019). *On the use of r and r squared in correlation and regression* (Vol. 34, No. 1, pp. 235-236). Hoboken, USA: John Wiley & Sons, Inc..
74. Ke, Y., Tang, X., & Jing, F. (2006, June). The design of high-level features for photo quality assessment. In *2006 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'06)* (Vol. 1, pp. 419-426). IEEE.
75. Kivy, P. (1991). Is music an art?.
76. Kong, S., Shen, X., Lin, Z., Mech, R., & Fowlkes, C. (2016, October). Photo aesthetics ranking network with attributes and content adaptation. In *European Conference on Computer Vision* (pp. 662-679). Springer, Cham.
77. Kruskal, W., & Wallis, W. (1952). Use of ranks in one-criterion variance analysis. *Journal of the American statistical Association*, 47(260), 583-621.
78. Kuhn, M., & Johnson, K. (2013). *Applied predictive modeling*. New York: Springer.
79. Lacassin, F. (1971). Pour un 9 ème art, la bande dessinée. *Union Générale d'Éditions, coll, 10*, 18.
80. Lander, D. (Ed.). (1990). *Sound by artists*. Art Metropole.

Autora: Iria María Santos López

Titor: Juan Jesús Romero Cardalda

Directores: Adrián Carballal Mato e Juan Jesús Romero Cardalda

Universidade da Coruña

81. Lansdown, J. (1988). Graphics, design and artificial intelligence. In *Theoretical Foundations of Computer Graphics and CAD* (pp. 1153-1174). Springer, Berlin, Heidelberg.
82. Leaver, T., Highfield, T., & Abidin, C. (2020). *Instagram: Visual social media cultures*. John Wiley & Sons.
83. Lecoutre, A., Negrevergne, B., & Yger, F. (2017, November). Recognizing art style automatically in painting with deep learning. In *Asian conference on machine learning* (pp. 327-342). PMLR.
84. Leder, H., Belke, B., Oeberst, A., & Augustin, D. (2004). A model of aesthetic appreciation and aesthetic judgments. *British journal of psychology*, 95(4), 489-508.
85. Lee, S., & Koubek, R. (2012). Users' perception of usability and aesthetics as criteria of pre-and post-use preferences. *European Journal of Industrial Engineering*, 6(1), 87-117.
86. Lempel, A., & Ziv, J. (1976). On the complexity of finite sequences. *IEEE Transactions on information theory*, 22(1), 75-81.
87. Letelier, J. (1937). Artes aplicadas. *Revista de Arte*, 3(14), 15-19.
88. Lewis, M. (2008). Evolutionary visual art and design. In *The art of artificial evolution* (pp. 3-37). Springer, Berlin, Heidelberg.
89. Li, C., & Chen, T. (2009). Aesthetic visual quality assessment of paintings. *IEEE Journal of selected topics in Signal Processing*, 3(2), 236-252.
90. Li, Y., Pu, Y., Xu, D., Qian, W., & Wang, L. (2017, October). Image aesthetic quality evaluation using convolution neural network embedded fine-tune. In *CCF Chinese Conference on Computer Vision* (pp. 269-283). Springer, Singapore.
91. Li, H., Zheng, Y., Wu, X., & Cai, Q. (2019). 3D model generation and reconstruction using conditional generative adversarial network. *International Journal of Computational Intelligence Systems*, 12(2), 697-705.
92. Liu, C. H., & Ting, C. K. (2016). Computational intelligence in music composition: A survey. *IEEE Transactions on Emerging Topics in Computational Intelligence*, 1(1), 2-15.
93. Liu, L., Chen, R., Wolf, L., & Cohen-Or, D. (2010, May). Optimizing photo composition. In *Computer Graphics Forum* (Vol. 29, No. 2, pp. 469-478). Oxford, UK: Blackwell Publishing Ltd.
94. Liu, S. C., & Chang, S. (1997). Dimension estimation of discrete-time fractional Brownian motion with applications to image texture classification. *IEEE Transactions on Image Processing*, 6(8), 1176-1184.
95. López de Mántaras, R. (2016). Artificial intelligence and the arts: Toward computational creativity.
96. Lukac, R. (Ed.). (2017). *Computational photography: methods and applications*. CRC press.

Autora: Iria María Santos López

Titor: Juan Jesús Romero Cardalda

Directores: Adrián Carballal Mato e Juan Jesús Romero Cardalda

Universidade da Coruña

97. Luo, Y., & Tang, X. (2008, October). Photo and video quality evaluation: Focusing on the subject. In *European Conference on Computer Vision* (pp. 386-399). Springer, Berlin, Heidelberg.
98. Machado, P., & Cardoso, A. (1998, November). Computing aesthetics. In *Brazilian Symposium on Artificial Intelligence* (pp. 219-228). Springer, Berlin, Heidelberg.
99. Machado, P., Romero, J., & Manaris, B. (2008). Experiments in computational aesthetics. In *The art of artificial evolution* (pp. 381-415). Springer, Berlin, Heidelberg.
100. Machado, P., Romero, J., Nadal, M., Santos, A., Correia, J., & Carballal, A. (2015). Computerized measures of visual complexity. *Acta psychologica*, 160, 43-57.
101. Maltese, C. (1980). *Las técnicas artísticas*. Cátedra.
102. Malu, G., Bapi, R. S., & Indurkha, B. (2017). Learning photography aesthetics with deep cnns. *arXiv preprint arXiv:1707.03981*.
103. Mann, H. B., & Whitney, D. R. (1947). On a test of whether one of two random variables is stochastically larger than the other. *The annals of mathematical statistics*, 50-60.
104. Mao, H., Cheung, M., & She, J. (2017, October). Deepart: Learning joint representations of visual arts. In *Proceedings of the 25th ACM international conference on Multimedia* (pp. 1183-1191).
105. Mapes, G. (2021). *Elite authenticity: Remaking distinction in food discourse*. Oxford University Press.
106. Marin, M. M., & Leder, H. (2013). Examining complexity across domains: relating subjective and objective measures of affective environmental scenes, paintings and music. *PloS one*, 8(8), e72412.
107. Mateas, M. (2003). Expressive AI: Games and Artificial Intelligence. *DiGRA Conference*.
108. McCormack, J., & d'Inverno, M. (2014). On the future of computers and creativity. *AISB 2014 Symposium on Computational Creativity*. London.
109. McWhinnie, H. J. (1968). A review of research on aesthetic measure. *Acta Psychologica*, 28, 363-375.
110. Meng, X., Gao, F., Shi, S., Zhu, S., & Zhu, J. (2018). Mlans: image aesthetic assessment via multi-layer aggregation networks. *2018 Eighth International Conference on Image Processing Theory, Tools and Applications (IPTA)* (pp. 1-6). IEEE.
111. Mensink, T., & Van Gemert, J. (2014, April). The rijksmuseum challenge: Museum-centered visual recognition. In *Proceedings of International Conference on Multimedia Retrieval* (pp. 451-454).
112. Murray, N., Marchesotti, L., & Perronnin, F. (2012, June). AVA: A large-scale database for aesthetic visual analysis. In *2012 IEEE Conference on Computer Vision and Pattern Recognition* (pp. 2408-2415). IEEE.

Autora: Iria María Santos López

Titor: Juan Jesús Romero Cardalda

Directores: Adrián Carballal Mato e Juan Jesús Romero Cardalda

Universidade da Coruña

113. Nadal, M., & Chatterjee, A. (2019). Neuroaesthetics and art's diversity and universality. *Wiley Interdisciplinary Reviews: Cognitive Science*, 10(3), e1487.
114. Nadikattu, R. R. (2014). Content analysis of American & indian Comics on Instagram using Machine learning. *International Journal of Creative Research Thoughts (UJCRT)*, 2320-2882.
115. Neumann, L., Sbert, M., Gooch, B., & Purgathofer, W. (2005). Defining computational aesthetics. *Computational aesthetics in graphics, visualization and imaging*, 13-18.
116. Oliveira, H. (2009). *Automatic generation of poetry: an overview*. Universidade de Coimbra.
117. Oliveira, H. G. (2012). PoeTryMe: a versatile platform for poetry generation. *Computational Creativity, Concept Invention, and General Intelligence*, 1, 21.
118. Parra Jaramillo, M. (2018). PERSPECTIVAS SOCIOLÓGICAS SOBRE LAS ARTES LITERARIAS. In *Ontología del arte: Tendencias filosóficas sobre el arte contemporáneo* (pp. 139-170). Universidad Pontificia Bolivariana.
119. Parra, F. (2017). *Estadística y Machine Learning con R*. RPubs Blog.
120. Pazos-Perez, R., Carballal, A., Rabuñal, J., Mures, O., & García-Vidaurrázaga, M. (2018). Predicting Vertical Urban Growth Using Genetic Evolutionary Algorithms in Tokyo's Minato Ward. *J. Urban Plan.*
121. Pena, M. L. C., Carballal, A., Rodríguez-Fernández, N., Santos, I., & Romero, J. (2021). Artificial intelligence applied to conceptual design. A review of its use in architecture. *Automation in Construction*, 124, 103550.
122. Penousal, F. J., & Santos, M. (2007). Computación Evolutiva y Creatividad. In *Inteligencia Artificial y Computación Avanzada* (p. 121).
123. Polatkan, G., Jafarpour, S., Brasoveanu, A., Hughes, S., & Daubechies, I. (2009, November). Detection of forgery in paintings using supervised learning. In *2009 16th IEEE International Conference on Image Processing (ICIP)* (pp. 2921-2924). IEEE.
124. Prasad, M., Jwala Lakshamma, B., Chandana, A., Komali, K., Manoja, M., Rajesh Kumar, P., & Sasi Kiran, P. (2018). An efficient classification of flower images with convolutional neural networks. *International Journal of Engineering & Technology*, 384-391.
125. Prevsner, N. (1982). *Las Academias de Arte*. Madrid: Cátedra.
126. Reinecke, K., Yeh, T., Miratrix, L., Mardiko, R., Zhao, Y., Liu, J., & Gajos, K. (2013). Predicting users' first impressions of website aesthetics with a quantification of perceived visual complexity and colorfulness. *Proceedings of the SIGCHI conference on human factors in computing systems*, (pp. 2049-2058).
127. Restrepo, L. F., & González, J. (2007). From pearson to Spearman. *Revista Colombiana de Ciencias Pecuarias*, 20(2), 183-192.

Autora: Iria María Santos López

Titor: Juan Jesús Romero Cardalda

Directores: Adrián Carballal Mato e Juan Jesús Romero Cardalda

Universidade da Coruña

128. Revert Gomis, J. (2016). Orígenes y fundamentos de la intermedialidad entre cine y cómic: una aproximación para la era digital.
129. Ritcher, G. (1959). *El arte griego*. Barcelona.
130. Roads, C. (1980). Artificial intelligence and music. *Computer Music Journal*, 4(2), 13-25.
131. Romero, J., & Machado, P. (2008). *The art of artificial evolution: A handbook on evolutionary art and music*. Springer Science & Business Media.
132. Romero, J., Machado, P., Carballal, A., & Correia, J. (2012a). Computing aesthetics with image judgement systems. In *Computers and creativity* (pp. 295-322). Springer, Berlin, Heidelberg.
133. Romero, J., Machado, P., Carballal, A., & Osorio, O. (2011, April). Aesthetic classification and sorting based on image compression. In *European Conference on the Applications of Evolutionary Computation* (pp. 394-403). Springer, Berlin, Heidelberg.
134. Romero, J., Machado, P., Carballal, A., & Santos, A. (2012b). Using complexity estimates in aesthetic image classification. *Journal of Mathematics and the Arts*, 6(2-3), 125-136.
135. Rosenblatt, F. (1958). The perceptron: a probabilistic model for information storage and organization in the brain. *Psychological review*, 65(6), 386.
136. Rovira, P. (2015). La construcción de las artes aplicadas en la arquitectura modernista: protagonismo de los oficios, los gremios y los artesanos en la ejecución material de obras arquitectónicas modernistas. In *Actas del Noveno Congreso Nacional y Primer Congreso Internacional Hispanoamericano de Historia de la Construcción: Segovia, 13 a 17 de octubre de 2015* (pp. 1520-1529). Instituto Juan de Herrera.
137. Sandoval, C., Pirogova, E., & Lech, M. (2019). Two-stage deep learning approach to the classification of fine-art paintings. *IEEE Access*, 7, 41770-41781.
138. Santos, I., Castro, L., Rodriguez-Fernandez, N., Torrente-Patino, A., & Carballal, A. (2021). Artificial Neural Networks and Deep Learning in the Visual Arts: A review. *Neural Computing and Applications*, 1-37.
139. Shiner, L. (2004). La invención del arte. In *Una historia cultural*. (EH Julibert, Trad.). Barcelona: Paidós.
140. Smith, J. M., & Maynard, S. J. (1993). *The theory of evolution*. Cambridge University Press.
141. Smuts, A. (2005). Are Video Games Art?. *Contemporary Aesthetics (Journal)*, 3(1), 6.
142. Snodgrass, J. G., & Vanderwart, M. (1980). A standardized set of 260 pictures: norms for name agreement, image agreement, familiarity, and visual complexity. *Journal of experimental psychology: Human learning and memory*, 6(2), 174.
143. Souriau, E. (1949). Time in the plastic arts. *The Journal of aesthetics and art criticism*, 7(4), 294-307.

Autora: Iria María Santos López

Titor: Juan Jesús Romero Cardalda

Directores: Adrián Carballal Mato e Juan Jesús Romero Cardalda

Universidade da Coruña

144. Spratt, E. L., & Elgammal, A. (2014, September). Computational beauty: Aesthetic judgment at the intersection of art and science. In *European Conference on Computer Vision* (pp. 35-53). Springer, Cham.
145. Stiny, G., & Gips, J. (1971, August). Shape grammars and the generative specification of painting and sculpture. In *IFIP Congress (2)* (Vol. 2, No. 3, pp. 125-135).
146. Talebi, H., & Milanfar, P. (2018). NIMA: Neural image assessment. *IEEE Transactions on Image Processing*, 27(8), 3998-4011.
147. Tan, W. R., Chan, C. S., Aguirre, H. E., & Tanaka, K. (2016a, September). Ceci n'est pas une pipe: A deep convolutional network for fine-art paintings classification. In *2016 IEEE international conference on image processing (ICIP)* (pp. 3703-3707). IEEE.
148. Tan, Y., Tang, P., Zhou, Y., Luo, W., Kang, Y., & Li, G. (2017). Photograph aesthetical evaluation and classification with deep convolutional neural networks. *Neurocomputing*, 228, 165-175.
149. Tan, Y., Zhou, Y., Li, G., & Huang, A. (2016b). Computational aesthetics of photos quality assessment based on improved artificial neural network combined with an autoencoder technique. *Neurocomputing*, 188, 50-62.
150. Taylor, R. P., Micolich, A. P., & Jonas, D. (1999). Fractal analysis of Pollock's drip paintings. *Nature*, 399(6735), 422-422.
151. Thielsch, M. T., Haines, R., & Flacke, L. (2019). Experimental investigation on the effects of website aesthetics on user performance in different virtual tasks. *PeerJ*, 7, e6516.
152. Thomas, M., Mainguy, C., & Pommier, S. (1985). *L'art textile*.
153. Thomsen Marcial, F. (1995). *Arte & diseño: las artes plásticas aplicadas al diseño gráfico* (Doctoral dissertation, Universidad del Salvador).
154. Tian, X., Dong, Z., Yang, K., & Mei, T. (2015). Query-dependent aesthetic model with deep learning for photo quality assessment. *IEEE Transactions on Multimedia*, 17(11), 2035-2048.
155. Tian, C., Xu, Y., Li, Z., Zuo, W., Fei, L., & Liu, H. (2020). Attention-guided CNN for image denoising. *Neural Networks*, 124, 117-129.
156. Toivonen, H., & Gross, O. (2015). Data mining and machine learning in computational creativity. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, 5(6) pp. 265-275.
157. Turing, A. M. (1988). Chess. *Computer Chess Compendium*, 14-17.
158. Upadhyaya, M., Dixit, D., & Pradesh, A. (2016). A review: Relating low level features to high level semantic in cbir. *International Journal of Signal Processing, Image Processing and Pattern Recognition*, 9(3), 433-444.
159. Valentine, C. W. (2005). The experimental psychology of beauty.

Autora: Iria María Santos López

Titor: Juan Jesús Romero Cardalda

Directores: Adrián Carballal Mato e Juan Jesús Romero Cardalda

Universidade da Coruña

160. Van Geert, E., & Wagemans, J. (2020). Order, complexity, and aesthetic appreciation. *Psychology of Aesthetics, Creativity, and the Arts*, 14(2), 135.
161. Viadel, R. M. (1997). «Enseñanza y aprendizaje en Bellas Artes: una revisión de los cuatro modelos históricos desde una perspectiva contemporánea». *Arte, individuo y sociedad*, (9), 55.
162. Wadjiz, A. Filsafat Estetika. (1980) Widagdo. Pluralisme Dalam Estetika. Orasi Ilmiah pada Sidang Terbuka Senat ITB. Penerbit Nur Cahya.
163. Wagner, M., Lin, H., Li, S., & Saupe, D. (2019). Algorithm selection for image quality assessment. *arXiv preprint arXiv:1908.06911*.
164. Wang, H., & Rasheed, K. (2014). Artificial intelligence in clothing fashion. In *Proceedings on the International Conference on Artificial Intelligence (ICAI)* (p. 1). The Steering Committee of The World Congress in Computer Science, Computer Engineering and Applied Computing (WorldComp).
165. Wang, Q., Luo, T., & Wang, D. (2016a, November). Can machine generate traditional chinese poetry? a feigenbaum test. In *International Conference on Brain Inspired Cognitive Systems* (pp. 34-46). Springer, Cham.
166. Wang, W., Cai, D., Wang, L., Huang, Q., Xu, X., & Li, X. (2016b). Synthesized computational aesthetic evaluation of photos. *Neurocomputing*, 172, 244-252.
167. Wang, Y., Huang, X., Lee, C. S., Zhang, S., Li, Z., Samaras, D., ... & Huang, P. (2004, September). High resolution acquisition, learning and transfer of dynamic 3-D facial expressions. In *Computer Graphics Forum* (Vol. 23, No. 3, pp. 677-686). Oxford, UK and Boston, USA: Blackwell Publishing, Inc.
168. Westlake, N., Cai, H., & Hall, P. (2016, October). Detecting people in artwork with cnns. In *European Conference on Computer Vision* (pp. 825-841). Springer, Cham.
169. White, C., & Matulionyte, R. (2019). *Artificial Intelligence Painting The Bigger Picture For Copyright Ownership*.
170. Woods, W., & Boudreau, J. (1950). Design complexity as a determiner of visual attention among artists and non-artists. *Journal of Applied Psychology*, 34(5), 355.
171. Wong, L. K., & Low, K. L. (2009, November). Saliency-enhanced image aesthetics class prediction. In *2009 16th IEEE International Conference on Image Processing (ICIP)* (pp. 997-1000). IEEE.
172. Wu, S. (2020, April). Development of Graphic Design Based on Artificial Intelligence. In *Journal of Physics: Conference Series* (Vol. 1533, No. 3, p. 032022). IOP Publishing.
173. Xia, Y., Liu, Z., Yan, Y., Chen, Y., Zhang, L., & Zimmermann, R. (2017). Media quality assessment by perceptual gaze-shift patterns discovery. *IEEE Transactions on Multimedia*, 19(8), 1811-1820.
174. Yannakakis, G., & Togelius, J. (2018). *Artificial intelligence and games*. New York: Springer.

Autora: Iria María Santos López

Titor: Juan Jesús Romero Cardalda

Directores: Adrián Carballal Mato e Juan Jesús Romero Cardalda

Universidade da Coruña

175. Yoshimura, Y., Cai, B., Wang, Z., & Ratti, C. (2019). Deep learning architect: Classification for architectural design through the eye of artificial intelligence. *International Conference on Computers in Urban Planning and Urban Management* (pp. 249-265). Springer.
176. Young-Min, K. (2018). What makes the difference in visual styles of comics: From classification to style transfer. *2918 3rd International Conference on Computational Intelligence and Applications (ICCIA)* (pp. 181-185). IEEE.
177. Young-Min, K. (2019). Feature visualization in comic artist classification using deep neural networks. *Journal of Big Data*, 6(1), 56.
178. Zeitlin, F. I., & Winkler, J. (1990). *Nothing to do with Dionysos*. Athenian drama in its social context.
179. Zhang, W., Ma, K., Yan, J., Deng, D., & Wang, Z. (2018). Blind image quality assessment using a deep bilinear convolutional neural network. *IEEE Transactions on Circuits and Systems for Video Technology*, 30(1), 36-47.
180. Zujovic, J., Gandy, L., Friedman, S., Pardo, B., & Pappas, T. N. (2009, October). Classifying paintings by artistic genre: An analysis of features & classifiers. In *2009 IEEE International Workshop on Multimedia Signal Processing* (pp. 1-5). IEEE.

ANEXOS (Publicacións)

Artigo 1: Santos, I., Castro, L., Rodríguez-Fernandez, N., Torrente-Patiño, Á., Carballal, A. *Artificial Neural Networks and Deep Learning in the Visual Arts: a review*. *Neural Comput & Applic* (2021). <https://doi.org/10.1007/s00521-020-05565-4>

[Q1; JCR ranking 2019: COMPUTER SCIENCE, ARTIFICIAL INTELLIGENCE 23/137; Impact Factor 4.774].

Artigo 2: Carballal, A., Cedron, F., Santos, I., Santos A., Romero J., *Minimal neural network topology optimization for aesthetic classification*. *Neural Comput & Applic* (2021). <https://doi.org/10.1007/s00521-020-05550-x>

[Q1; JCR ranking 2019: COMPUTER SCIENCE, ARTIFICIAL INTELLIGENCE 23/137; Impact Factor 4.774].

Artigo 3: Carballal, A., Fernandez-Lozano, C., Rodriguez-Fernandez, N., Santos, I., & Romero, J. (2020). *Comparison of Outlier-Tolerant Models for Measuring Visual Complexity*. *Entropy*, 22(4), 488. <https://doi.org/10.3390/e22040488>

[Q2; JCR ranking 2019: MATHEMATICS, INTERDISCIPLINARY APPLICATIONS 28/106; MULTIDISCIPLINARY SCIENCES 31/71; Impact Factor 2.462].