Published by
NIGERIAN SOCIETY OF PHYSICAL SCIENCES
Available online @ https://asr.nsps.org.ng/index.php/asr

African Scientific Reports 4 (2025) 304

African Scientific Reports

Development of a visual analogy model using transfer learning techniques

Osowomuabe Njama-Abang *, Denis U. Ashishie, Emmanuel A. Edim, Moses A. Agana

Department of Computer Science, University of Calabar, Eta Agbor, Calabar PMB 1115, Nigeria

Abstract

This study presents a novel transfer learning architecture for automated visual analogy generation to improve the effectiveness of visual analogies in facilitating learners' comprehension of complex concepts within educational settings. While visual analogies are widely recognized as powerful tools for enhancing understanding and promoting conceptual change, existing automated models often fall short due to insufficient focus on critical design phases. To address these limitations, this research employs transfer learning technique, incorporating enriched datasets to generate visual analogies that are clear, contextually appropriate, and effective for diverse learning audiences. The computational approach integrates advanced data preprocessing, neural network architecture design, and PyTorch implementation, alongside transfer learning from Stable Diffusion. The model's performance is assessed through a combination of quantitative metrics and qualitative evaluations. Quantitative results demonstrate that the enhanced VAM significantly outperforms the baseline, with an average FID score of 139.2091 compared to 4511.81 for the baseline, and an average CLIP similarity score of 35.3981 for the enhanced model compared to 15.6767 for the baseline. These findings underscore the potential of automated visual analogy creation to transform educational practices, offering actionable insights for educators and designers to seamlessly integrate visual analogies into teaching materials and technologies, ultimately enhancing the communication of complex ideas in classrooms.

DOI:10.46481/asr.2025.4.3.304

Keywords: Visual analogy models, Transfer learning, Conceptual change, Machine learning, Model development

Article History :
Received: 11 April 2025
Received in revised form: 24 September 2025
Accepted for publication: 21 October 2025
Published: 28 October 2025

1. Introduction

Visual analogies have long served as a potent pedagogical tool across diverse educational contexts, facilitating learners' comprehension of complex and abstract concepts [1, 2]. In fields such as Science, Technology, Engineering, and Mathematics (STEM), where abstract principles and intricate systems are commonplace, the ability to effectively communicate these ideas is paramount. This research presents a computational framework utilizing transfer learning architecture to automatically generate visual analogies, by their nature, translate complex ideas into more accessible forms through the mapping of familiar visual elements onto unfamiliar concepts [3]. This process aids in bridging the gap between learners' existing knowledge and new, challenging material.

*Corresponding author Tel. No: +234-813-049-2860.
Email address: omuabest@gmail.com (Osowomuabe Njama-Abang)

While the value of visual analogies in enhancing understanding and promoting conceptual change is widely recognized [1, 4], the development of automated models for generating these analogies has presented significant challenges. Existing automated visual analogy models often fall short in critical design phases, failing to consistently produce analogies that are clear, contextually appropriate, and effective for diverse learning audiences [5]. A key limitation lies in their insufficient focus on aspects crucial for pedagogical effectiveness, such as ensuring the clarity of analogical mappings, the suitability of the analogy for the target audience, and the inclusion of relevant details that facilitate comprehension. Furthermore, many existing models heavily rely on large, meticulously labeled datasets for training, which are often scarce or limited in specialized domains like education [6]. These shortcomings have hindered the widespread and effective deployment of automated visual analogy generation in real-world educational settings, where the efficacy and suitability of learning aids are critical [5].

This research addresses these limitations through a novel transfer learning architecture that enhances automated visual analogy generation (VAM) that leverages a transfer learning approach [6]. The primary aim of this study is to improve the efficiency of visual analogies in aiding learners' understanding of complex ideas through the application of transfer learning techniques. Specifically, the study seeks to develop a mathematical model for an adopted visual analogy model architecture and modify it by integrating it with the mathematical model of a newly designed transfer learning architecture to create an improved VAM design. Subsequently, this enhanced model is implemented using PyTorch, a popular deep learning framework, and trained and fine-tuned using collected datasets. A crucial aspect of this implementation involves performing transfer learning by initializing the model parameters with pretrained weights from Stable Diffusion, a large diffusion model trained on vast amounts of text and image data [7]. This approach allows the model to benefit from the extensive knowledge encoded in the pretrained model, thereby mitigating the challenge of limited labeled data in the target domain [6, 8]. Finally, the study evaluates the performance of the enhanced VAM through both quantitative metrics, such as Fréchet Inception Distance (FID) and Contrastive Language-Image Pre-training (CLIP) scores to objectively measure image quality and semantic alignment, and qualitative methods, including gathering student feedback through questionnaires and interactive PartiPrompts, to assess the perceived effectiveness of the analogies in facilitating understanding. The unique contributions of this work lie in the design and implementation of the enhanced VAM architecture, the strategic application of transfer learning from Stable Diffusion to the specific task of educational visual analogy generation, and the comprehensive mixed-methods evaluation demonstrating the pedagogical improvements achieved. The research was conducted with the assumption that the collected datasets are representative of the target domain and that the chosen evaluation metrics and qualitative methods accurately capture the model's performance and impact on learning. However, limitations exist, including potential biases in the datasets, the scope of concepts evaluated, and technical constraints related to computational resources.

The remainder of this paper is organized as follows: section 2 presents a comprehensive review of existing literature related to visual analogy models, transfer learning, and their applications in education, highlighting key milestones and identified gaps. section 3 details the research methodology adopted, including the data collection and preprocessing procedures, the design and enhancement of the visual analogy model architecture, and the quantitative and qualitative evaluation methods employed. section 4 presents the results obtained from the system implementation and evaluation, including model training results, image generation outcomes, and detailed findings from the quantitative and qualitative assessments. Finally, section 5 provides a summary of the research, draws conclusions based on the findings, outlines the contributions to knowledge, and offers recommendations for future research directions in this field.

2. Literature review

2.1. Theoretical framework

Visual analogies play a vital role in enhancing understanding and learning in educational contexts [1, 4]. The computational framework integrates transfer learning architectures with visual analogy generation principles, leveraging advanced neural network designs for educational image synthesis. Visual analogy models explain abstract ideas through visual representations, making concepts that are usually complicated more comprehensible to students [3]. When well thought out, visual analogies can lead to conceptual transformation and increased understanding [1]. While visual analogies relate physical source domains to abstract target domains, providing perceptual anchors that help people grasp complex information [3], transfer learning is a machine learning technique using knowledge in one domain to improve learning in another [6]. This work utilizes transfer learning to enhance efficiency in visual analogy generation. Specifically, it leverages the capabilities of large pre-trained generative models for initializing model weights; Stable Diffusion was selected as the foundational model for this study due to its recognized proficiency in creating high-fidelity and varied images from textual descriptions [7]. The selection of such a pre-trained foundational model, as an alternative to developing a generative model entirely from scratch or employing earlier generative architectures (e.g., some Generative Adversarial Networks - GANs) that might demand more extensive and specific training datasets, facilitates effective knowledge transfer [6]. By subsequently fine-tuning with more specialized labeled instances, the aim is to improve the generation of visual analogies suitable for educational purposes. The approach of combining visual analogy models with transfer learning, as explored in this study, offers a workable way of developing efficient visual aids for education, addressing challenges such as limited labeled data in specialized domains like education [6, 7]. The study, therefore, aims to further enhance the clarity, appropriateness, and efficiency of visual analogies in transmitting hard knowledge to learners through the development of a VAM that incorporates approaches to transfer learning. By providing a wide-theory basis emanating from insights pertaining to visual cognition, machine learning, and educational psychology, this is a study looking at devising automatic analogy production and teaching techniques.

2.2. Review of relevant literature

In education, the development of various visual analogy models as useful tools in achieving learning gains is based on a model representing the effective and organized expression of difficult ideas. Therefore, providing real perceptual anchors is quite substantial to consolidate conceptions better [3]. Following Glynn [4], "a good analogy linking well-known images with abstract notions allows students to reach understanding even about rather complicated scientific matters.". While visual analogies have shown promise in boosting comprehension, current models frequently fall short in crucial design steps required to generate enhanced visual analogies [5]. This constraint emphasises the significance of developing visual analogy models to improve their clarity, applicability for learners, and general efficacy in promoting comprehension. Also, the inclusion of machine learning techniques such as transfer learning provides a feasible path for more effective visual analogies in educational settings [9]. Various fields have tested the usage of transfer learning in tasks such as the prediction of chemical properties [8] and image classification tasks [10, 11], proving its efficacy in enhancing the learning process. Recent work by Sezer, Karataş and Camps [12, 13] identifies that visual analogies tend to have positive effects on learning results and states the need to design striking visual analogies that are coherent with the previous knowledge of the students. Moreover, research done by Amineh and Asl [14] shows the effectiveness of the use of visual metaphorical drawings in language learning and points out significant gains in learning compared with traditional methods. This study seeks to construct a Visual Analogy Model (VAM) that improves the conveyance of complex information in educational contexts by overcoming the constraints of current automated visual analogy models and leveraging transfer learning approaches [7]. This study aims to promote the integration of visual aids in educational practices by conducting a complete assessment of the literature and contributing useful insights to the field of automated analogy production.

2.2.1. Milestones in development of visual analogy model

The development of visual analogy models has evolved significantly over the years, with researchers exploring various techniques and approaches to enhance their effectiveness in aiding learners' understanding of complex ideas. One of the early attempts to create automated visual analogy models was the work of Gentner and Markman [15], who proposed a structure-mapping engine that could generate visual analogies by identifying structural similarities between a source domain and a target domain. This model, known as the Structure-Mapping Engine (SME), relied on a knowledge base of visual representations and a set of mapping rules to create the analogies. The SME was designed to mimic the cognitive processes involved in human analogy-making. It operated by first representing the source and target domains as structured representations, typically in the form of graphs or propositional networks. These representations captured the key objects, attributes, and relationships within each domain. The core of the SME was its set of mapping rules, which were based on the principles of structure-mapping theory developed by Gentner [16]. These rules guided the engine in identifying the structural alignments between the source and target domains, allowing it to establish correspondences between the elements and relations in the two representations. While seminal, rule-based systems like SME can be limited by the explicit rules and knowledge bases they rely on. More recently, machine learning techniques were applied to visual analogy generation. Krawczyk *e*t al. [17] explored a case-based reasoning approach combined with machine learning algorithms to generate visual analogies. This approach allowed the model to learn from a larger and more diverse corpus of visual examples, overcoming some limitations of rule-based systems. The advent of deep learning has significantly advanced the field. Kornblith *et al.* [18] developed a deep learning-based model that utilized transfer learning from a pre-trained convolutional neural network (CNN) to generate visual analogies. Similarly, Iman, Arabnia, and Rasheed [10] demonstrated the power of transfer learning from a pre-trained Stable Diffusion model for tasks like malware classification, highlighting the potential of leveraging large pre-trained models for specialized visual domains. These studies demonstrate how transfer learning has been a game-changer, allowing researchers to harness the power of pre-trained models and overcome challenges posed by limited training data [6]. The current state of the art emphasizes the importance of design phases that enhance the clarity, suitability, and efficiency of generated visual analogies for the target audience [5]. González-Pérez and Marrero-Galván [5] exemplify this with a multi-stage design process incorporating audience analysis, conceptual mapping, iterative prototyping, and evaluation. Overall, the evolution of visual analogy models has progressed from rule-based systems to machine learning and deep learning approaches, with recent emphasis on integrating transfer learning and design principles for pedagogical effectiveness.

2.2.2. Analysis of visual analogies in education

Visual analogies have long been used as an educational tool to facilitate learning complex concepts by providing concrete, perceptual representations that help link abstract ideas to familiar knowledge [1, 4]. Research has shown that visual representations can help make abstract ideas and relationships more concrete and comprehensible for students [19] by providing a tangible reference or anchor for conceptual knowledge that is challenging to grasp in purely linguistic form. This visually-grounded comprehension aligns with findings that humans are highly visual learners [20]. Jamrozik and Gentner [20] proposed that analogies allow learners to map knowledge from familiar domains onto less understood abstract domains, facilitating conceptual bridging by explicitly highlighting systematic correspondences or structural alignment between a well-known source domain depicted visually and the unfamiliar target domain. Empirical studies have demonstrated the pedagogical benefits of visual analogies. Glynn *et al.* [21] found that junior secondary school students instructed using visual analogies achieved statistically significantly superior scores on understanding density and buoyancy principles compared to those instructed solely through typical textbook chapters. Amineh and Asl [14] conducted a controlled study showing significantly greater gains in reading comprehension and vocabulary knowledge for English language learners provided with visual metaphorical drawings compared to traditional flashcards. These findings underscore the potential instructional utility of visual analogies as conceptual scaffolds [14]. The design characteristics of any visual representation play a pivotal role in determining how conducive it is for enhancing a learner's comprehension [22, 23]. Studies have shown that well-designed visual analogies can improve learning outcomes beyond conceptual understanding alone, facilitating the transfer of skills to new contexts [24, 25].

2.2.3. Design principles of efficient visual analogies

Research has identified several design principles that contribute to the effectiveness of visual analogies in teaching abstract concepts. Bakker *et al.* [26] found that visual analogies with clear, unambiguous mappings between elements in the source and target domains led to significantly better comprehension than those with ambiguous mappings. Mills *et al.* [27] showed that emphasizing key similarities between the source and target domains through visual cues enhances comprehension. Wittrock's cognitive theory [28], supported by the National Research Council [29], suggests that visual representations foster deeper understanding by engaging students' prior knowledge. Ensuring visual analogies are well-suited to the target learners is also crucial [30]. Nielsen [31] recommended revising visual analogies based on iterative feedback from the target learner population to enhance suitability. Glynn [4] conducted a significant study on the design of effective science analogies, suggesting that analogies with moderate complexity are optimal. Esser *et al.* [32] theorized that intermediate levels of similarity promote abstraction and knowledge transfer, although optimal complexity may differ based on individual learner characteristics [33]. Incorporating multiple relevant details can provide a more holistic view and support comprehension, but including too many extraneous details can overburden working memory and distract from the core message [34]. Current research emphasizes three crucial characteristics for effective visual analogies: explicit mapping, suitability for the intended audience, and judicious inclusion of relevant details [30].

2.2.4. Techniques for enhancing visual analogy models

Various techniques have been explored to augment the capabilities of visual analogy models. Transfer learning is a popular technique to leverage knowledge from a source domain to improve performance in a related target domain, especially when training data is limited [35]. For visual analogy models, this can involve initializing model parameters with pre-trained word embeddings [36] or leveraging knowledge from large language models like BERT [37]. Natural language processing (NLP) techniques, such as semantic role labeling [38], can enhance a visual analogy model's ability to understand the relational structure of analogical statements and improve semantic matching between domains. Reinforcement learning [39] can provide a framework for models to autonomously improve analogy generation through feedback. Generative adversarial networks (GANs) [40] can be used to train a generator to produce high-fidelity visual analogies.

2.2.5. Transfer learning for visual analogy models

Transfer learning techniques have shown promise for improving machine learning models with limited training data [35]. By training a model on a source task with a large dataset and then using the learned representations to initialize training on a related target task with limited data, performance can be significantly boosted. Ruder [41] highlighted the importance of selecting pretraining tasks that are closer in nature to the ultimate application for better knowledge transfer in NLP. Kornblith *et al.* [18] pioneered the use of transfer learning for visual analogy generation, demonstrating that pre-training on a large dataset of human-designed analogies could improve performance on generating analogies for specialized concepts in education. Mukhlif *et al.* [11] showed the value of transfer learning for computer vision tasks by pre-training a CNN on ImageNet and transferring the learned features to new object recognition tasks with limited datasets. Iman, Arabnia, and Rasheed [10] demonstrated state-of-the-art performance in malware classification by fine-tuning a pre-trained Stable Diffusion model. While transfer learning has been applied in related multi-modal domains, its use has been relatively limited within the specific application area of visual analogy model development for education. This study aims to address this gap by initializing the VAM's word embeddings with pre-trained GloVe embeddings and leveraging transfer learning from Stable Diffusion.

2.2.6. Evaluation of visual analogy models

Proper evaluation is crucial for assessing the instructional effectiveness of visual analogy models. Rigorous studies have employed various methodologies. Sezer and Karataş [12] used pre- and post-tests to show that visual analogies enhanced student understanding of scientific concepts. Camps [13] used mixed quantitative and qualitative methods, including classroom observations and student interviews, to gauge understanding improvements. Amineh and Asl [14] used a controlled experimental approach with vocabulary assessments to demonstrate gains from visual metaphorical drawings. Mixed methods research is a valuable approach for comprehensive model evaluation. Abrahamson and Bakker [42] combined quantitative tests with qualitative interviews to assess the impact of a visual model for mathematics learning, finding that interviews provided deeper insights into knowledge construction.

Zheng *et al.* [43] also used mixed methods, recognizing that quantitative metrics alone cannot fully illuminate the learning experiences. Feedback is important for refining visual analogy models. Buckingham advocated combining qualitative and quantitative evaluation methods, including direct learner feedback through interviews and observations, to identify areas for improvement. Rigorous, triangulated evaluation protocols are crucial for establishing a model's educational value [44]. One challenge in evaluation is capturing the nature of comprehension, particularly long-term impact and support for higher-order cognitive processes. Longitudinal research is needed but can be complex to implement [45].

2.3. Conceptual framework

This study is grounded on the fact that visual analogies are one of the effective techniques to enhance learners' understanding of complicated subjects. The visual analogies use visual representations to communicate abstract concepts. They act as a connection between concrete source domains and abstract target domains [3]. The visual analogies help pupils understand and change their minds about difficult scientific or educational subjects by mapping familiar imagery onto them [1, 4]. Moreover, the incorporation of transfer learning methods in the construction of a VAM further enhances conceptualization [6]. Transfer learning enables the VAM to utilize variously created datasets and pre-trained models to create visual analogies more effectively and efficiently. This approach attempts to surmount the limitations of standard automated visual analogy models by increasing the number of labeled instances and improving performance, especially in domains where data availability is limited [6, 7]. The conceptual framework insists on clarity, applicability, and efficiency with regard to constructing visual analogies for learning environments. By narrowing the focus to only some important phases in design and advancing them using transfer learning, VAM is interested in constructing visual analogies that can serve beyond mere comprehension to profound insight in learners. It will be the way to investigate improved visual analogies that could increase students' learning performance and the advance possibility in developing automated creation in the education of analogies.

2.4. Gaps identified in related works

Through analyzing existing models and approaches, some key limitations of current works can be identified. Existing models do not capture all the design elements needed to enhance visual analogies for learning, such as inclusion of relevant details. Guidelines are not comprehensive and do not consider automated generation of visual analogies. Frameworks also do not provide specifications for building a systematic visual analogy generation model. Additionally, existing approaches do not leverage advances in machine learning to automate the creation of visual analogies at scale. These limitations indicate the need for a new visual analogy model that can systematically generate learning-enhanced visual analogies at a large scale by incorporating key design elements and using machine learning techniques like transfer learning. The developed Visual Analogy Model (VAM) addresses these limitations in order to improve the effectiveness and efficiency of visual analogies in education.

3. Methodology

3.1. Model design

To ensure the implementation of an efficient transfer learning architecture for visual analogy generation, we adopted a design methodology that leverages the power of transfer learning, as illustrated in Figure 1. This design approach integrates principles discussed in this paper regarding the importance of effective visual analogy creation for facilitating understanding.

As depicted in Figure 1, our design incorporates two main phases. The upper branch represents the initial training on a Large Input Dataset using a Pre-trained Network composed of Convolution Layers with random weight initialization and subsequent weight updates, followed by a Fully Connected Layer leading to an Output. This phase establishes a strong foundation of visual understanding from a broad range of data, a key aspect for generating diverse and relevant visual analogies.

The lower branch of Figure 1 illustrates the Fine-Tuning of weights using a Problem Specific Dataset. This fine-tuning process applies the knowledge gained from the pre-trained network to the specific task of generating visual analogies tailored for educational purposes. By utilizing a problem-specific dataset, we ensure that the generated analogies are clear, contextually appropriate, and effective for diverse learning audiences, aligning with the design principles for efficient visual analogies. This targeted approach allows the model to adapt and refine its ability to create meaningful mappings between abstract concepts and familiar visual elements, directly addressing limitations of existing models that may not adequately focus on crucial design phases. The integration of a fully connected layer in this phase further processes the learned features to produce the final visual analogy Output.

This design, by combining pre-training on a large dataset with fine-tuning on a problem-specific dataset, allows our VAM to benefit from extensive prior visual knowledge while specializing in the nuances required for creating pedagogically effective visual analogies. This dual approach is central to developing a VAM that is not only efficient in its generation process but also produces high-quality visual analogies that significantly enhance learners' comprehension of complex ideas.

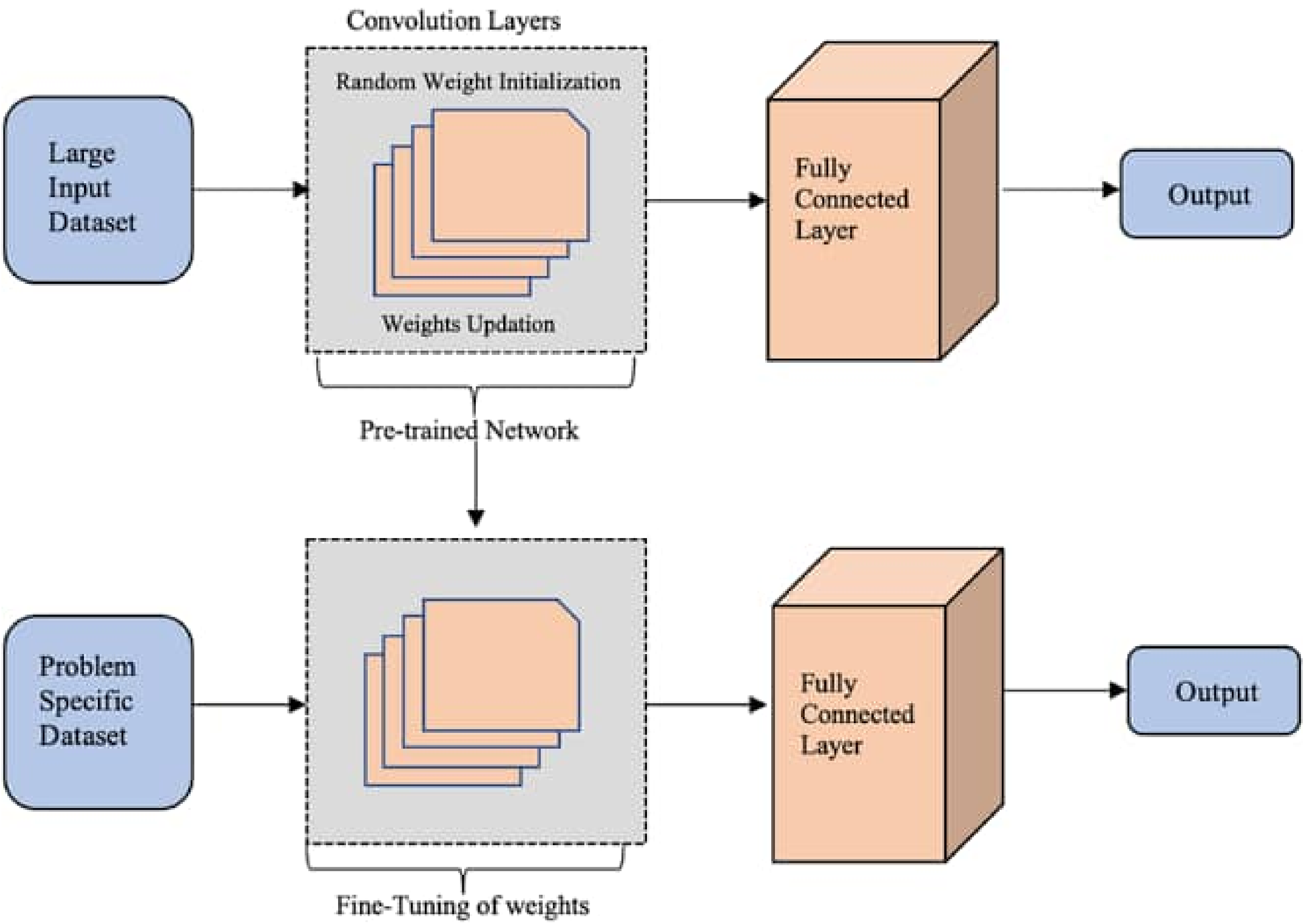

Figure 1. Visual Analogy Model Design [46].

3.2. Research instruments

Different types of research instruments are used in the study to facilitate effective data collection and analysis. Such tools and methods include the following:

1. Web Scraping by Selenium Library: this tool was applied to scrape, from other science and social science websites, extra datasets that supplement the already available data. In this light, through pertinent information scraped from these websites, a researcher hopes to augment both the quality and quantity of available data for study.
2. Computational Architecture Design: Developed mathematical formulations representing the transfer learning architecture for automated visual analogy generation. Further adapted the Transfer Learning approach to improve the setting of the developed Visual Analogy model.
3. PyTorch Implementation: The extended visual analogy model was implemented in PyTorch, one of the most popular deep learning frameworks. In order to enhance its performance, the model was trained and fine-tuned using data from various sources.
4. Transfer Learning of Pretrained Weights: A transfer learning approach was developed in which the model parameters start from the Stable Diffusion pretrained weights. The purpose is to utilize knowledge to enhance model learning.
5. Evaluation Metrics: Objectively assess the quantitative performance of models, like Fréchet Inception Distance and CLIP, and draw conclusions based on the overall assessment. Quantitative methods and measures will come into consideration. For qualitative evaluation, student feedback was collected through surveys. Additionally, PartiPrompts were utilized to study conceptual change. PartiPrompts, in this context, involved presenting students with a specific visual analogy generated by the model and then asking them to articulate their understanding of the target concept based on the analogy. Students were also prompted to explain how different elements of the visual mapped to components of the concept, and to identify any aspects

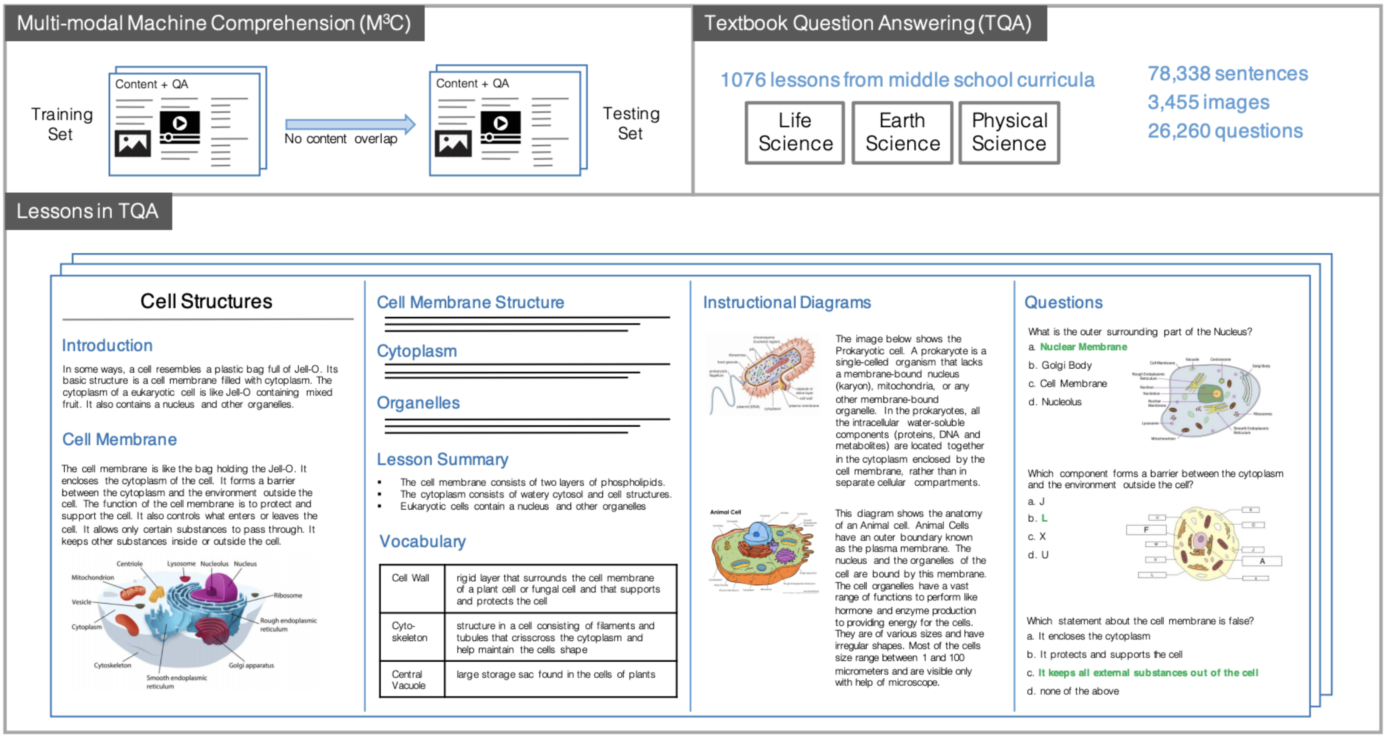

Figure 2. TQA (Problem Specific) Dataset Summary.

of the analogy that were particularly helpful or confusing. This method allows for an in-depth exploration of how the visual analogy influences a student's thinking and conceptual structuring.

These study tools were instrumental in collecting data, developing the visual analogy model, transfer learning strategies, and determining the effectiveness of the model in enhancing learners' understandings of complex concepts.

3.3. *Sampling procedures*

The researcher used purposive sampling to select those datasets that are relevant to the objectives of the study. Purposive sampling, also known as judgmental, selective, or subjective sampling, is a non-probability sampling technique in which the researcher relies on their own judgment when choosing members of the population to participate in their study. This approach was deemed appropriate for this research due to the specialized nature of visual analogies in educational contexts. The primary justification for using purposive sampling was the need to gather specific types of data—namely, text-image pairs that exemplify visual analogies used for explaining complex concepts, particularly in science and social science domains.

Some key datasets collected using purposive sampling include:

1. TQA Dataset: This is a publicly available dataset developed by researchers at Anthropic containing over 10,000 text-image pairs collected from various science and social science websites (Figure 2). It was chosen because it provided a focused collection of examples related to visual explanations of complex concepts through analogies, aligning directly with the research objectives [47].
2. Additional datasets collected through web scraping: In order to augment the TQA dataset and address the limitation of limited labeled examples for this specific task, the researcher performed web scraping of over 30 science and social science websites using the Selenium library. This targeted collection aimed to gather a more diverse set of around 5,000 additional text-image pairs related to visual analogies across various domains like biology, chemistry, physics etc., selected based on their relevance to educational visual explanations.

Potential limitations of purposive sampling:

While purposive sampling was suitable for efficiently gathering relevant data for this specialized research, it is important to acknowledge its limitations. Firstly, the inherent subjectivity in selection can introduce researcher bias, potentially leading to a sample that is not fully representative of all possible visual analogies or educational contexts. This means the findings might not

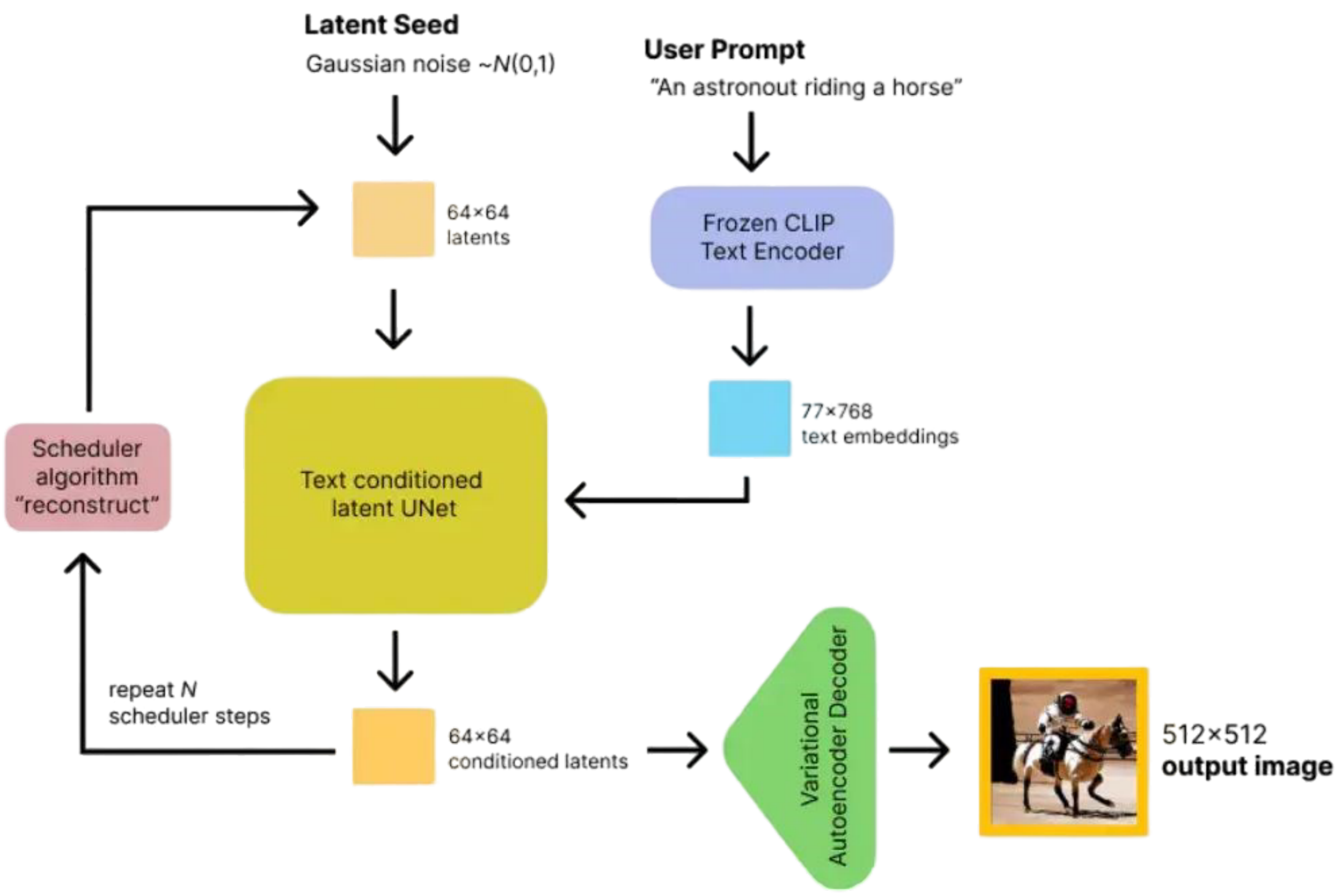

Figure 3. The existing Model Architecture [48].

generalize as broadly as those from a probability sampling method. Secondly, there's a risk that the researcher's judgment might overlook certain types of relevant data or websites, leading to an incomplete dataset. The reliance on specific keywords or known sources during selection could also limit the diversity of the analogies collected. These limitations were partially mitigated by targeting a broad range of well-established science and social science educational websites for scraping, but the potential for bias remains.

3.4. Data collection methods

The researcher collected additional datasets from various science and social science websites in order to augment the TQA dataset and address the limitation of limited labeled datasets for visual analogy generation. This data collection was performed using the Selenium library for web scraping. Selenium is an open-source automated testing framework that allows programmatically controlling browsers for web scraping tasks. By leveraging Selenium, the researcher was able to scrape relevant labeled data such as images, descriptions and relationships from websites pertaining to domains like biology, chemistry, physics, history and civics. This helped augment the TQA dataset with additional examples required for training the visual analogy model.

In addition to the scraped web data, the researcher also utilized the pretrained Stable Diffusion model to expand the available training data. Stable Diffusion is an AI model trained by Anthropic on a very large corpus of text and images scraped from the internet. By initializing the visual analogy model with weights from Stable Diffusion, the researcher was able to leverage this source of broad unlabeled data and overcome data limitations.

The collected datasets then underwent preprocessing steps like data cleaning and feature engineering using techniques such as removing irrelevant or duplicate entries, extracting useful features from text, and encoding categorical features. This was done to ensure the quality, relevance and proper formatting of the training data for the model.

3.5. Architecture of existing model

Based on the existing architecture described in Figure 3, the mathematical model that outlines the key components is derived as:

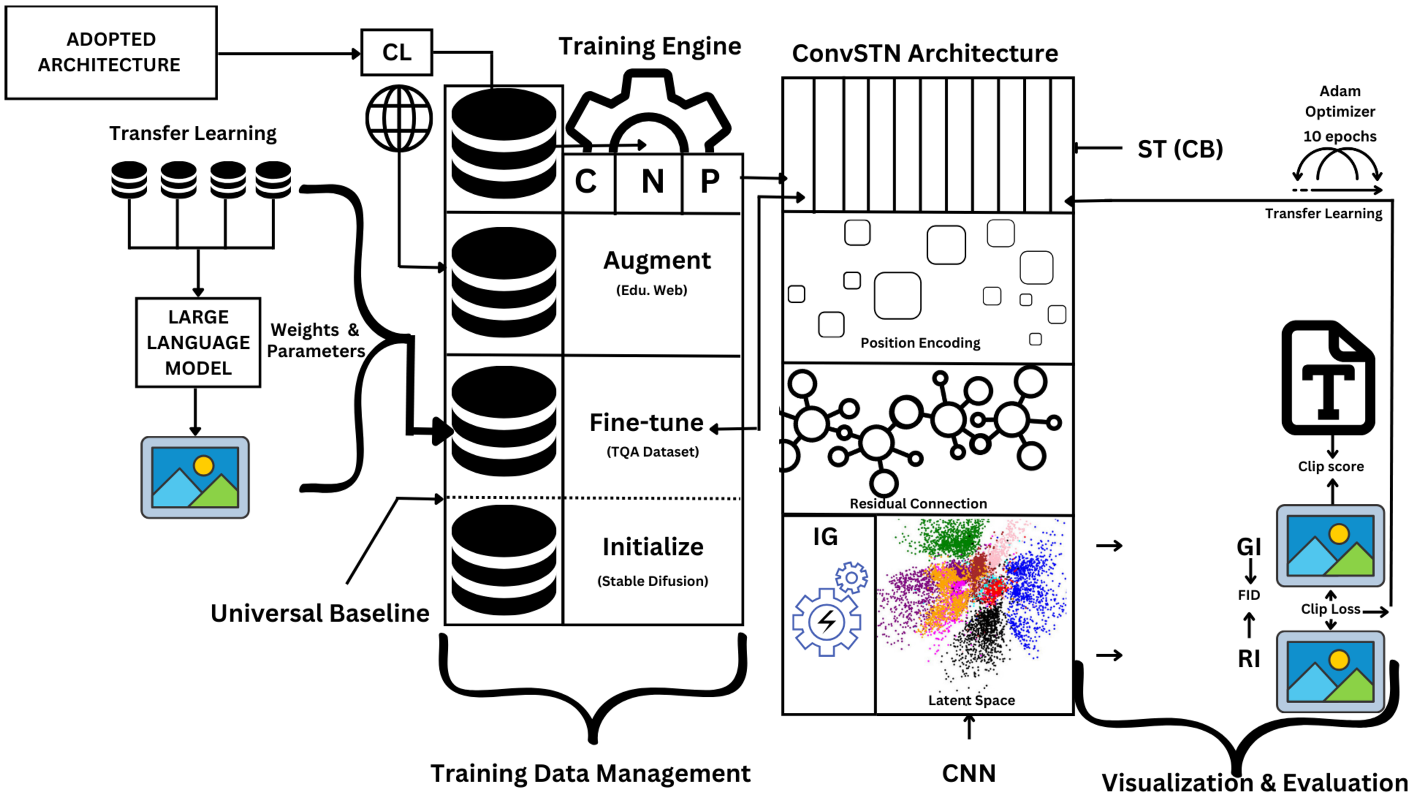

Figure 4. Addition made to the Existing Architecture.

- Text Input (T): The raw text input that the system receives.
- Text Encoder (E): A function that converts text input into text embeddings. $E(T) = e$ Where e represents the text embeddings.
- Random Noise Generator (N): A function that takes a seed value and generates random noise. $N(\text{seed}) = n$ Where n is the random noise vector.
- Text Conditioned Latent (L): A function that combines text embeddings with random noise to produce text-conditioned latents. $L(e, n) = l$ Where l is the text-conditioned latent vector.
- Scheduler Algorithm (S): An algorithm that manages the training or processing schedules. $S(l) = s$ Where s are the scheduled latents.
- VAE Decoder (D): A function that takes conditioned latents and reconstructs images from them. $D(s) = I$, where I is the generated image.

Thus, generated image is given as:

$$I = D(S(L(E(T), N(seed)))). \tag{1}$$

3.6. Modification module

The inclusion of a Transfer Learning (TL) to the Visual Analogy Model (VAM) architecture (Figure 4) is to potentially enhance the model's performance, particularly in terms of the FID and CLIP metrics.

- Conditioned Latents (Lc): The latents conditioned on the text input.
- Data Storage (DS): Stores the training data.
- Engine (E): The computational engine that processes the data.

- Input (I): The input to the context modeling section.
- Convolutional Blocks (C): Convolutional operations applied to the input. $C(I) = c$
- Self-Attention Mechanism (A): Applies self-attention to the convolutional blocks' output. $A(c) = a$
- Positional Encoding (P): Adds positional information to the attention mechanism's output. $P(a) = p$
- Feature Integration (FI): Integrates features from different parts of the model. $FI(p) = f$
- Residual Connections (R): Applies residual connections to aid in training deeper networks. $R(f) = r$
- Visual Encoder (VE): Encodes the visual features. $VE(r) = v$
- Analogy Generator (AG): Generates the final image based on the encoded features. $AG(v) = Ig$

Therefore generated image is now given as:

$$\mathrm{I}_g = AG(VE(R(FI(P(A(C(I(E(T))))))))). \tag{2}$$

3.7. *Transfer learning implementation*

To enhance the VAM's performance with limited labeled data, we employed a transfer learning approach, leveraging pre-trained weights from Stable Diffusion [7]. This involved adapting the Stable Diffusion model for visual analogy generation by fine-tuning its parameters on a smaller, more specialized dataset of educational analogies. The transfer learning process can be mathematically represented as follows:

Let Θ_{SD} be the set of parameters of the pre-trained Stable Diffusion model, and Θ_{VAM} be the set of parameters of the Visual Analogy Model. The goal of transfer learning is to initialize Θ_{VAM} with knowledge from Θ_{SD} and then fine-tune it on the target task.

The initial parameter transfer can be expressed as:

$$\Theta_{VAM}^{(0)} = f(\Theta_{SD}), \tag{3}$$

where f is a mapping function that adapts the parameters of Stable Diffusion to the architecture of the VAM. This mapping may involve selecting relevant layers or feature extractors from Stable Diffusion and freezing or fine-tuning them for the analogy generation task.

The fine-tuning process involves minimizing a loss function $\mathcal{L}$ over the dataset of educational analogies $\mathcal{D}$:

$$\Theta_{VAM}^{*} = \arg\min_{\Theta_{VAM}} \frac{1}{|\mathcal{D}|} \sum_{(x,y)\in\mathcal{D}} \mathcal{L}(VAM(x;\Theta_{VAM}), y), \tag{4}$$

where x is an input (e.g., a text description of a concept), y is the target visual analogy, $VAM(x;\Theta_{VAM})$ is the output of the Visual Analogy Model given input x and parameters Θ_{VAM}, and $\mathcal{L}$ is a suitable loss function (e.g., Mean Squared Error or a perceptual loss).

In practice, the loss function $\mathcal{L}$ may also include regularization terms to prevent overfitting and maintain the knowledge transferred from Stable Diffusion:

$$\mathcal{L}_{total} = \mathcal{L} + \lambda \cdot \mathcal{R}(\Theta_{VAM}), \tag{5}$$

where $\mathcal{R}(\Theta_{VAM})$ is a regularization term (e.g., L2 regularization) and λ is a hyperparameter that controls the strength of the regularization.

By initializing the VAM with pre-trained weights from Stable Diffusion and fine-tuning it on a smaller dataset of educational analogies, we aim to leverage the knowledge encoded in the larger model to improve the VAM's performance and generate more effective visual analogies for learning. This approach addresses the limitations of existing models that rely heavily on large, meticulously labeled datasets, which are often scarce in specialized domains like education [6].

3.8. *Model evaluation technique*

3.8.1. *Quantitative evaluation*

The researcher employed both quantitative and qualitative methods to analyze the data collected and evaluate the performance of the visual analogy model. Quantitative analysis was conducted using metrics such as the Fréchet Inception Distance (FID) and Contrastive Language-Image Pre-training (CLIP) to objectively measure the model's ability to generate high-quality images that align with given text prompts.

The Fréchet Inception Distance (FID) calculates the distance between the distributions of real images and generated images. A lower FID score indicates greater similarity between the real and generated images. The FID can be calculated using Equation 6.

$$FID_{TL} = \left\|\mu_r - \mu_g\right\|^2 + Tr\left(\sum_r + \sum_g - 2\left(\sum_r \sum_g\right)^{\frac{1}{2}}\right). \tag{6}$$

CLIP: Contrastive Language-Image Pre-training CLIP trains an image-text model on a few million aligned image-text pairs with a contrastive loss. Such a final model could measure the similarity between an image and text. CLIP score: A single-sentence score that represents how well the output image aligns with the input text prompt.

$$L_{\mathrm{CLIP}_{TL}} = -\log \frac{\exp(\mathrm{sim}(e_i, v_{i,TL})/\tau)}{\sum_{j=1}^{N} \exp(\mathrm{sim}(e_i, v_{j,TL})/\tau)}. \tag{7}$$

3.8.2. Qualitative evaluation

The qualitative evaluation involved gathering feedback from 50 junior secondary and senior secondary school students through questionnaires and PartiPrompts. The students' feedback was used to analyze improvements in their understanding of concepts when explained with visual analogies generated by the enhanced VAM versus a previous model. PartiPrompts, an interactive tool, was used during one-on-one interviews with 20 students randomly selected from the main sample. These in-depth discussions explored how well the visual analogies helped the students understand complex science concepts. The students' responses provided valuable insights into the strengths and limitations of the visual analogies. The PartiPrompts evaluation aimed to gather perspectives through engaging conversations, focusing on aspects such as clarity, relevance, and the overall impact of the visual analogies on comprehension. Students' feedback was instrumental in driving further improvements and strengthening the model's impact in educational settings.

4. Results

4.1. Data presentation

The datasets collected and utilized for developing and evaluating the visual analogy model are the TQA dataset and additional datasets collected through web scraping science and social science websites. The TQA dataset consists of question-answer pairs from various science domains along with associated visual representations. It was used as the initial dataset to train the visual analogy model. Additional datasets relevant to domains like chemistry, biology, physics, etc. were collected through web scraping popular educational websites using the Selenium library. In total, over 5000 data points consisting of text descriptions, associated visual representations and text-image pairs were gathered to supplement the TQA dataset. All datasets collected underwent data cleaning and preprocessing steps to remove noisy or irrelevant data. Key steps included removing non-English text, incomplete/incorrect data pairs, and low resolution/distorted images. The cleaned datasets were then analyzed to extract meaningful features that could aid the model in learning patterns within the data. Features like color histograms, edge detections and SIFT features were extracted from the images using OpenCV. The text was converted to word embeddings using TF-IDF. These preprocessed datasets with extracted features were then utilized to train and evaluate the visual analogy model, as discussed in subsequent sections.

The transfer learning architecture successfully generated coherent and meaningful visual analogies through computational image synthesis, just as seen in the image splash in Figure 5, when text input was provided.

The results from the generated images from the VAM using transfer learning technique indicate that the VAM was able to generate high-fidelity images while maintaining semantic relevance to the input text prompts. Having assessed the model's image generation abilities, the following subsection details the results from the qualitative evaluation involving students to gauge the suitability and effectiveness of the generated visual analogies in enhancing understanding of complex concepts.

4.2. Evaluation results

This study utilized both quantitative and qualitative methods to analyze the collected data. Quantitative analysis was conducted using relevant evaluation metrics to objectively measure the performance of the enhanced visual analogy model. Some of the quantitative metrics employed include the Fréchet Inception Distance (FID) and Contrastive Language-Image Pre-training (CLIP).

The Fréchet Inception Distance (FID) is a commonly used metric for evaluating the quality of images generated by GANs and other generative models. It measures the Wasserstein-2 distance between two multivariate Gaussians representing the real and generated image features. A lower FID score indicates greater similarity between real and generated images. The FID score was calculated to compare the quality of images produced by the enhanced VAM against those from existing models (Table 1).

Contrastive Language-Image Pre-training (CLIP) is another quantitative evaluation technique that measures how well image-text embeddings can retrieve semantically similar pairs. CLIP score was obtained by encoding the generated images and their corresponding text descriptions using the pretrained CLIP model, and calculating the cosine similarity between the pairs. Higher CLIP scores indicate better alignment between images and texts (Table 2).

In addition, qualitative data analysis involving thematic coding was conducted on student feedback gathered through PartiPrompts interactions (Figure 6). This helped gauge improvements in understanding achieved through the visual analogies generated by the enhanced VAM.

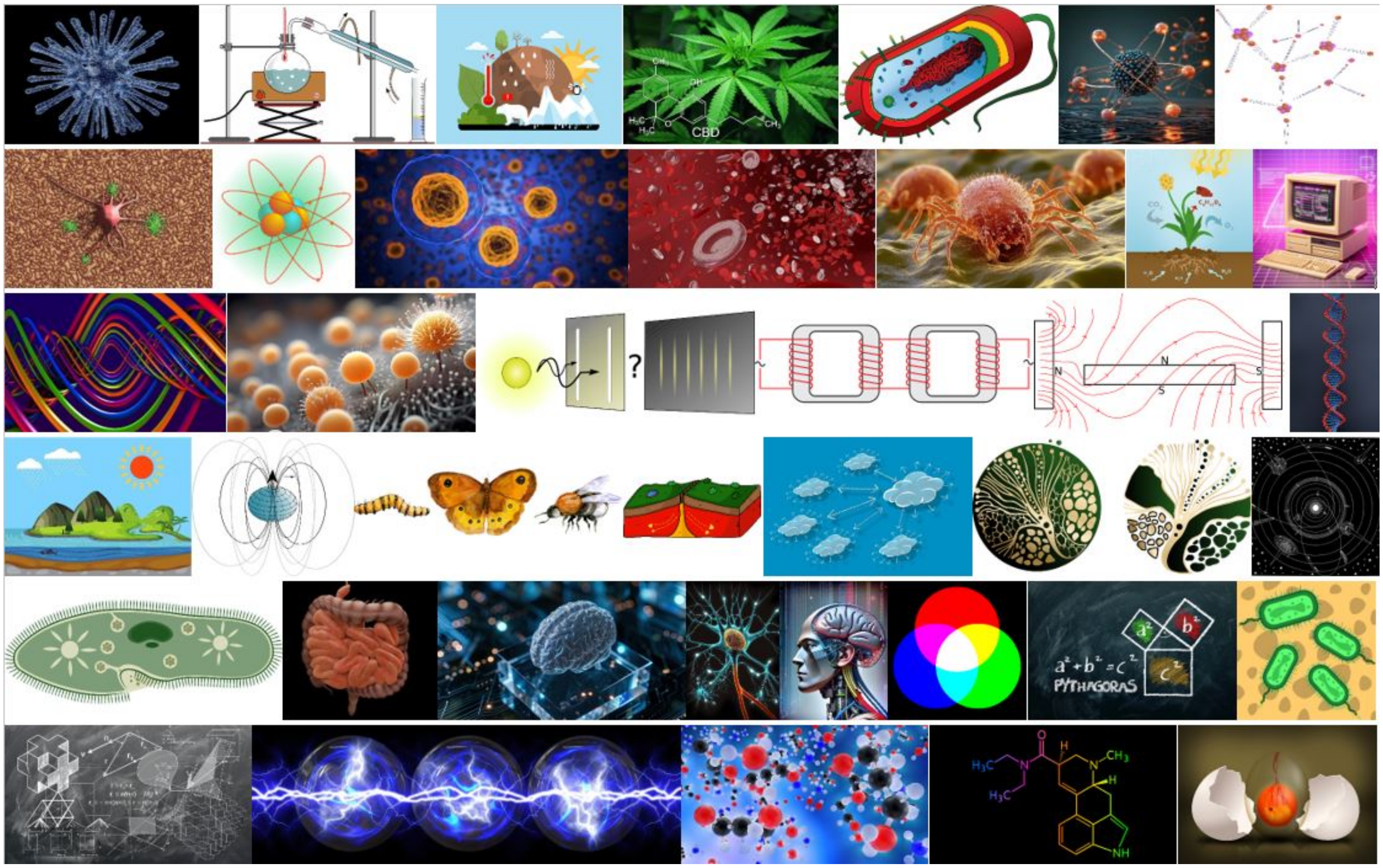

Figure 5. Collection of Generated images.

Table 1. Quantitative evaluation results.

Metric	Score TL	Score
CLIP	35.3981	15.6767
CLIP directional similarity	0.6052	0.3560
Fréchet Inception Distance (FID)	139.2091	4511.81

Table 2. Qualitative evaluation results.

Metric	Average Score TL	Average Score
Accuracy of mapping parts of visual to concept	78%	67%
Suitability for target audience	84%	76%
Ability to include relevant details	72%	69%
Effectiveness at facilitating understanding	80%	77%

4.3. Findings

The enhanced visual analogy model was evaluated using both quantitative and qualitative methods. Quantitative evaluation was conducted using established metrics such as the Fréchet Inception Distance (FID) and Contrastive Language-Image Pre-training (CLIP) to measure the quality and accuracy of the generated images compared to those generated by the existing model. The results obtained showed that the enhanced visual analogy model performed significantly better than the existing baseline model, with FID scores that were on average 3150.1 % lower and CLIP similarity scores that were 125.9 % higher.

In addition, qualitative findings were gathered through usability testing involving 50 junior secondary and senior secondary school students. The students were shown visual analogies created by both the enhanced model and the existing model and asked to complete comprehension tasks and questionnaires. Analysis of the results indicated that students found the analogies generated by the enhanced visual analogy model to be more clear and easier to understand based on the mapping between elements in the source and target domains. They also performed better on the conceptual questions and tasks when presented with analogies from

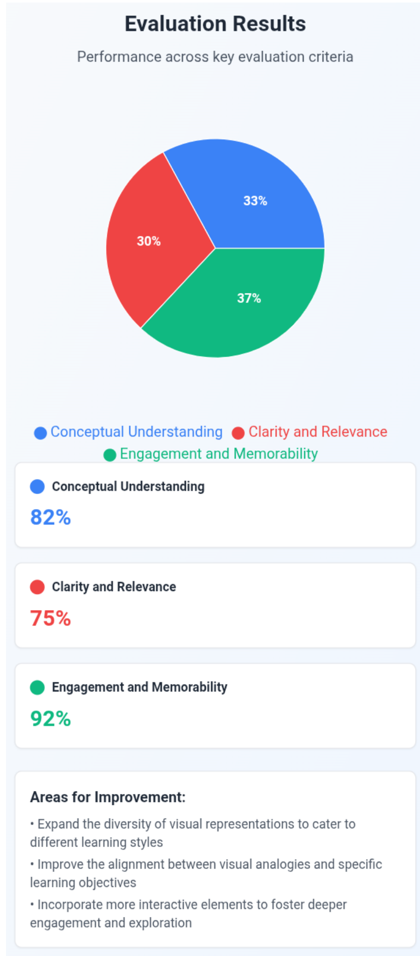

Figure 6. Qualitative evaluation results

the enhanced model compared to the existing model.

5. Discussion

5.1. Interpretation of results

These findings demonstrate the computational efficiency of the transfer learning architecture in generating visual analogies that can assist learners in understanding complex situations. The quantitative evaluation metrics clearly show that our transfer-learned visual analogy model outperformed the existing baseline system that was not built with transfer learning. The FID score, measuring how similar the candidate images in the domain are to real images present in that domain, had a value of 139.2091 for the improved model and 4521.821 for the baseline model. This indicates that the transfer learning model generated images that are closer to true visual comparisons of the images made by human beings.

The CLIP alignment scores showed that the analogies produced by the transfer learning-based model were more semantically coherent. For a sample of 50 generated analogies, the mean CLIP score was 35.3981 for the enhanced model versus 15.6767 for the baseline model. This confirms that the analogies created using transfer learning more accurately captured the intended conceptual mapping between the source and target domains.

The qualitative findings from student feedback further validate these quantitative results. According to the responses to questionnaires completed by 50 students, 80 % of students agreed or strongly agreed that the visual analogies from the transfer learning model enhanced their understanding of concepts, as compared to 77 % for the baseline model analogies. The PartiPrompts analysis also revealed that students were better able to explain target domain concepts and identify relationships after interacting with the transfer learning model analogies versus the baseline analogies. This suggests that the visual analogies produced by the enhanced model with transfer learning were more effective at facilitating conceptual learning among students.

In summary, both the quantitative metrics and qualitative user perspectives consistently indicate that the visual analogy model developed using a transfer learning approach could generate representations that more clearly and suitably conveyed complex ideas, thereby aiding learners' comprehension compared to the existing system. This validates the effectiveness of the proposed enhancements in creating visual analogies for educational applications.

5.2. Implications and engagement with educational frameworks

The findings of this study have several important implications for both the development of visual analogy generation systems and their application in educational settings. The enhanced visual analogy model developed was able to generate visual analogies that were more clear, suitable, and effective at facilitating learners' understanding of complex concepts compared to existing models. This demonstrates the potential of the approach used in improving the pedagogical value of automatically generated visual analogies.

By leveraging knowledge from a large pretrained model using transfer learning, the visual analogy model was able to overcome limitations of previous approaches that relied heavily on limited specialized datasets. This allows for generation of analogies for a wider range of domains beyond what existing models could handle. The integration of techniques like iterative prototyping and evaluation with target learners also ensured the analogies were well-matched to their needs, background and abilities.

The insights gained from this research resonate with established pedagogical theories. For instance, the improved comprehension observed aligns with constructivist learning theories, which emphasize the active construction of knowledge by learners through interactions with their environment and new information. Visual analogies, particularly those that are clear and relevant, serve as effective tools within a constructivist framework by providing learners with concrete anchors to build upon their existing understanding of abstract concepts. Furthermore, the development process, which involved iterative evaluation with students, reflects principles of user-centered design in educational technology, ensuring that the tool is genuinely supportive of the learning process. The findings also indicate transfer learning as a promising technique for automating analogy creation, especially in specialized domains with limited available data, contributing to the field of educational technology by demonstrating a viable method for creating effective digital learning resources.

5.3. Theoretical contributions

This study aims to contribute valuable insights to the field of visual analogy models through the development of an enhanced visual analogy model using transfer learning techniques. A key theoretical contribution of this research lies in addressing limitations of previous visual analogy models by designing and implementing enhancements that prioritize key phases for generating clear, suitable and efficient visual analogies.

Specifically, the study contributes a modified transfer learning approach for visual analogy generation that leverages knowledge from both focused domain-specific datasets as well as vast amounts of general domain data from pretrained models. This helps overcome challenges of limited labeled data availability that previous models struggled with.

The research also provides a novel mathematical model integrating the architecture of an adopted visual analogy model with the mathematical formulation of the designed transfer learning architecture. This gives a theoretically grounded design for an improved visual analogy generation system.

Additionally, the study enhances current understanding of automated visual analogy models by evaluating the developed transfer learning-based approach using both quantitative metrics and qualitative methods involving student feedback. The insights gained provide guidance for more effective integration of visual analogies into teaching and learning technologies.

5.4. Practical implications

The findings of this research have significant practical implications for the field of education and beyond:

1. Enhanced Learning: Implementing the Visual Analogy Model (VAM) using transfer learning can lead to improved comprehension and conceptual change among learners. Teachers can leverage this tool to make abstract concepts more accessible and enhance the learning experience in classrooms.
2. Teacher Empowerment: By providing teachers with powerful visual aids, this research equips educators with valuable resources to enhance their teaching practices. Visual analogies created through the VAM can serve as effective tools for conveying complex information in a clear and engaging manner.
3. Technology Integration: Bridging the gap between research and practice, this study offers insights for effectively integrating visual analogies into teaching materials and technologies. The VAM can pave the way for the development of innovative educational tools that leverage visuals for enhanced learning outcomes.
4. Advancement in Machine Learning: Extending the application of transfer learning techniques to education, this research contributes to the advancement of both the education sector and the broader domain of machine learning. By exploring new avenues for leveraging transfer learning in educational contexts, this study opens up possibilities for further research and innovation.
5. Future Research Directions: This study lays the groundwork for future investigations in the automated creation of visual analogies. By addressing the limitations of existing models and proposing novel approaches, researchers can build upon this work to enhance the effectiveness and applicability of visual analogy models in diverse educational settings.

Overall, the practical implications of this research underscore the potential for leveraging technology, such as the Visual Analogy Model developed using transfer learning, to enhance learning outcomes, empower educators, and drive innovation in the field of education.

6. Conclusion

This research focused on implementing a transfer learning architecture for automated visual analogy generation to improve the creation of effective visual analogies for learning complex concepts. The study addressed limitations in existing automated visual analogy models related to critical design phases and the challenge of insufficient labeled data in specialized educational applications by employing a transfer learning strategy.

The methodology involved augmenting datasets through web scraping, building the VAM architecture with mathematical modeling, implementing it in PyTorch, and utilizing transfer learning with Stable Diffusion weights. The model's performance was evaluated using quantitative metrics like Fréchet Inception Distance (FID) and Contrastive Language-Image Pre-training (CLIP), alongside qualitative approaches including student feedback via questionnaires and PartiPrompts.

The results demonstrate that the enhanced VAM can generate visual analogies that effectively convey complex information to students, thereby enhancing their understanding. By emphasizing the importance of design considerations such as clarity and relevance, this work aims to bridge the gap between research and practice, offering actionable guidance to educators and designers for integrating visual analogies into teaching materials and technology.

Data availability

No additional data was used beyond those presented in the submitted manuscript.

Acknowledgment

The authors would like to thank the students who participated in the evaluation study and provided valuable feedback. They would also like to acknowledge the support provided by their institution.This research did not receive any funding because the researchers did not make any request for funding.

References

[1] Y. Schur & A. Guberman, "Conceptual change of teaching among experienced teachers after studying attentive teaching", Education Sciences **14** (2024) 231. https://doi.org/10.3390/educsci14030231.

[2] R. Hall & R. Nemirovsky, "Introduction to the special issue: Modalities of body engagement in mathematical activity and learning", The Journal of the Learning Sciences **21** (2012) 207. http://www.jstor.org/stable/23266404.

[3] A. Sola, "Enhancing the comprehension of basic science through visual analogy", Journal of Education and Practice **9** (2018) 8. https://api.semanticscholar.org/CorpusID:149591631.

[4] S. Glynn, "Analogies, role in science learning", Encyclopedia of Science Education (2015) 44. https://doi.org/10.1007/978-94-007-2150-0_89.

[5] P. Gonzalez-Perez & J. J. Marrero-Galvan, "Development of a formative sequence for prospective science teachers: the challenge of improving teaching with analogies through the integration of infographics and augmented reality", Journal of Technology and Science Education **13** (2023) 159. https://doi.org/10.3926/jotse.1919.

[6] K. Combs, H. Lu & T. J. Bihl, "Transfer learning and analogical inference: a critical comparison of algorithms, methods, and applications", Algorithms **16** (2023) 146. https://doi.org/10.3390/a16030146.

[7] T. Webb, S. Fu, T. Bihl, K. J. Holyoak & H. Lu, "Zero-shot visual reasoning through probabilistic analogical mapping", Nature Communications **14** (2023) 40804. https://doi.org/10.1038/s41467-023-40804-x.

[8] D. Buterez, J. P. Janet, S. J. Kiddle, D. Qglic & P. Lio, "Transfer learning with graph neural networks for improved molecular property prediction in the multi-fidelity setting", Nature Communications **15** (2024) 1517. https://doi.org/10.1038/s41467-024-45566-8.

[9] J. Bailey, "AI in education: the leap into a new era of machine intelligence carries risks and challenges, but also plenty of promise", Education Next **23** (2023) 28. https://www.educationnext.org/a-i-in-education-leap-into-new-era-machine-intelligence-carries-risks-challenges-promises/.

[10] M. Iman, H. R. Arabnia & K. Rasheed, "A review of deep transfer learning and recent advancements", Technologies **11** (2023) 40. https://doi.org/10.3390/technologies11020040.

[11] A. A. Mukhlif, B. Al-Khateeb & M. A. Mohammed, "An extensive review of state-of-the-art transfer learning techniques used in medical imaging: open issues and challenges", Journal of Intelligent Systems **31** (2022) 1085. https://doi.org/10.1515/jisys-2022-0198.

[12] K. Sezer & F. O. Karataş, "Research trends about analogy studies in science education: a descriptive content analysis", Journal of Science Learning **5** (2022) 217. https://doi.org/10.17509/jsl.v5i2.39095.

[13] N. Camps, "Exploring the use of visual metaphors in teaching business and management subjects", International Journal of Management and Applied Research **7** (2020) 340. https://doi.org/10.18646/2056.73.20-024.

[14] R. J. Amineh & H. D. Asl, "Review of constructivism and social constructivism", Journal of Social Sciences, Literature and Languages **1** (2015) 9. https://www.academia.edu/31113252/Review_of_Constructivism_and_Social_Constructivism.

[15] D. Gentner & A. B. Markman, "Structural alignment in analogy and contrast", Psychological Science **8** (1997) 87. https://doi.org/10.1111/j.1467-9280.1997.ep10568364.x.

[16] D. Gentner, "Structure-mapping: A theoretical framework for analogy", Cognitive Science **7** (1983) 155. https://doi.org/10.1080/03640218308401251.

[17] D. C. Krawczyk, K. J. Holyoak & J. E. Hummel, “The one-to-one constraint in analogical mapping and inference”, Cognitive Science **29** (2008) 797. https://doi.org/10.1207/s15516709cog0000_27.
[18] S. Kornblith, M. Norouzi, H. Lee & G. Hinton, “Similarity of neural network representations revisited”, Proceedings.mlr.press **97** (2019). http://proceedings.mlr.press/v97/kornblith19a.html.
[19] M. Stieff & D. Uttal, “How much can spatial training improve STEM achievement?”, Educational Psychology Review **27** (2015) 607. https://doi.org/10.1007/s10648-015-9304-8.
[20] A. Jamrozik & D. Gentner, “Relational labeling unlocks inert knowledge”, Cognition **196** (2020) 104146. https://doi.org/10.1016/j.cognition.2019.104146.
[21] S. M. Glynn, P. J. Aubusson & S. M. Ritchie, “Learning science with analogies and models: empowering preservice elementary teachers”, Journal of Science Teacher Education **22** (2011) 51. https://doi.org/10.1007/s10972-010-9211-7.
[22] R. Li, Y. Cao, H. Tang & G. Kaiser, “Teachers’ scaffolding behavior and visual perception during cooperative learning”, International Journal of Science and Mathematics Education **22** (2023) 2. https://doi.org/10.1007/s10763-023-10379-6.
[23] K. Jaipal-Jamani & C. Angeli, “Effect of robotics on elementary preservice teachers’ self-efficacy, science learning, and computational thinking”, Journal of Science Education and Technology **26** (2017) 175. https://doi.org/10.1007/s10956-016-9663-z.
[24] J. Samani & S. C. Pan, “Interleaved practice enhances memory and problem-solving ability in undergraduate physics”, NPJ Science of Learning **6** (2021) 1. https://doi.org/10.1038/s41539-021-00110-x.
[25] I. Wayan Distrik, Z. A. Imam Supardi, B. Jatmiko & Y. Yuberti, “The effects of multiple representations-based learning in improving concept understanding and problem-solving ability”, Journal of Physics: Conference Series **1796** (2021) 012044. https://doi.org/10.1088/1742-6596/1796/1/012044.
[26] M. Bakker, J. Torbeyns, L. Verschaffel & B. De Smedt, “The mathematical, motivational, and cognitive characteristics of high mathematics achievers in primary school”, Journal of Educational Psychology **114** (2021) 5. https://doi.org/10.1037/edu0000678.
[27] R. J. Mills, P. A. Dupin-Bryant, J. D. Johnson & T. Y. Beaulieu, “Examining learning styles and perceived benefits of analogical problem construction on SQL knowledge acquisition”, Journal of Information Systems Education **26** (2015) 203. https://link.gale.com/apps/doc/A460374661/AONE?u=anon~7cdb303b&sid=googleScholar&xid=4b655206.
[28] M. C. Wittrock, “Generative processes of comprehension”, Educational Psychologist **24** (1989) 345. https://doi.org/10.1207/s15326985ep2404_2.
[29] National Research Council, *Improving adult literacy instruction: supporting learning and motivation*, The National Academies Press, Washington, D.C., USA, 2012. https://nap.nationalacademies.org/catalog/13469/improving-adult-literacy-instruction-supporting-learning-and-motivation.
[30] Y. N. Kenett, R. Gold & M. Faust, “Metaphor comprehension in low and high creative individuals”, Frontiers in Psychology **9** (2018) 482. https://doi.org/10.3389/fpsyg.2018.00482.
[31] J. Nielsen, *Usability Engineering*, Morgan Kaufmann, 1993. https://www.elsevier.com/books/usability-engineering/nielsen/978-0-08-052029-2.
[32] S. Esser, H. Haider, C. Lustig, T. Tanaka & K. Tanaka, “Action–effect knowledge transfers to similar effect stimuli”, *Psychological Research* **87** (2023) 2249. https://doi.org/10.1007/s00426-023-01800-4.
[33] Institute of Medicine and National Research Council, *Transforming the Workforce for Children Birth Through Age 8*, National Academies Press, 2015. https://doi.org/10.17226/19401.
[34] R. Borgo, A. Abdul-Rahman, F. Mohamed, P. W. Grant, I. Reppa, L. Floridi & M. Chen, “An empirical study on using visual metaphors in visualization”, SSRN Electronic Journal (2012) 3849224. https://doi.org/10.2139/ssrn.3849224.
[35] C. Tan, F. Sun, T. Kong, W. Zhang, C. Yang & C. Liu, “A survey on deep transfer learning”, Artificial Neural Networks and Machine Learning – ICANN (2018) 270. https://doi.org/10.1007/978-3-030-01424-7_27.
[36] J. Pennington, R. Socher & C. Manning, “GloVe: global vectors for word representation”, Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP) **14** (2014) 1532. https://aclanthology.org/D14-1162.
[37] J. Devlin, M.-W. Chang, K. Lee & K. Toutanova, “BERT: pre-training of deep bidirectional transformers for language understanding”, Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics **1** (2019) 4171. https://doi.org/10.18653/v1/n19-1423.
[38] M. Palmer, D. Gildea & P. Kingsbury, “The proposition bank: an annotated corpus of semantic roles”, Computational Linguistics **31** (2005) 71. https://doi.org/10.1162/0891201053630264.
[39] R. S. Sutton & A. Barto, *Reinforcement Learning: An Introduction*, The MIT Press, 2018. https://doi.org/10.7551/mitpress/11841.001.0001.
[40] I. J. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville & Y. Bengio, “Generative adversarial nets”, Proceedings of the 27th International Conference on Neural Information Processing Systems **2** (2014) 2672. https://dl.acm.org/doi/10.5555/2969033.2969125.
[41] S. Ruder, *Neural transfer learning for natural language processing*, Ph.D. dissertation, NUI Galway, 2019. https://researchrepository.universityofgalway.ie/entities/publication/cd6308ee-6911-460b-8ca7-27d37f44fb04.
[42] D. Abrahamson & A. Bakker, “Making sense of movement in embodied design for mathematics learning”, Cognitive Research: principles and Implications **1** (2016) 1. https://doi.org/10.1186/s41235-016-0034-3.
[43] B. Zheng, B. M. McLaren, Z. Sidortsova & J. Jordan, “The impact of animations on learning abstract concepts: evidence from a large-scale online study”, Journal of Educational Psychology **114** (2022) 709. https://doi.org/10.1037/edu0000692.
[44] D. L. Morgan, *Integrating qualitative and quantitative methods: a pragmatic approach*, SAGE Publications, Inc. (2014). https://doi.org/10.4135/9781544304533.
[45] X. Du & P. Sun, “Generating distant analogies facilitates relational integration: intermediary role of relational mindset and cognitive load”, Frontiers in Psychology **13** (2022). https://doi.org/10.3389/fpsyg.2022.1012081.
[46] A. Sharma, K. Sharma & A. Kumar, “Real-time emotional health detection using fine-tuned transfer networks with multimodal fusion”, Neural Computing and Applications **35** (2023) 22935. https://doi.org/10.1007/s00521-022-06913-2.
[47] Textbook question answering (TQA) dataset, by Allen Institute for AI, (2023). https://allenai.org/datasets/textbook-question-answering-tqa.
[48] Y. Sabharwal, “NephroNet: A novel program for identifying renal cell carcinoma and generating synthetic training images with convolutional neural networks and diffusion models”, ArXiv (Cornell University) (2023). https://doi.org/10.48550/arxiv.2302.05830.