JÖNKÖPING UNIVERSITY
School of Engineering

Dating of fashion plates (1820-1880) using transfer learning

Recognition of the year of origin of fashion plates

PAPER WITHIN *Computer Science, Artificial Intelligence*
AUTHORS: *Magdalena Dobrzańska*
TUTOR: *Florian Westphal*
JÖNKÖPING *June 2023*

This exam work has been carried out at the School of Engineering in Jönköping in the subject area Computer Science, Artificial Intelligence. The work is a part of the two-year Master of Science in Engineering programme. The authors take full responsibility for opinions, conclusions and findings presented.

Examiner:	Vladimir Tarasov
Supervisor:	Florian Westphal
Scope:	7,5 credits
Date:	2023-06-15

Mailing address:	Visiting address:	Phone:
Box 1026	Gjuterigatan 5	036-10 10 00 (vx)
551 11 Jönköping		

Contents

Abstract

Fashion history is an integral subfield of history as a whole. Fashion plates provide important evidence of what fashion once looked like and as such are a valuable window into the lives of the people in ages past. The rise of digitization opens new avenues for aiding historians in the dating of fashion plates; following on from this, digitization also brings a greater need for artworks to be digitized and AI can be utilized in order to keep up with the demand. This provides unique challenges such as gathering data and working with a relatively limited database. Due to the lack of prior research into the subject of dating fashion plates using Artificial Intelligence (AI), said application of AI in the dating process could help future historians automate the task. Transfer learning can help streamline the dating process of fashion plates. I used several approaches with three different models (ResNet101, NasNetMobile, and InceptionV3) and achieved the best mean absolute error of 2.8 years in a range of 60 years using NasNetMobile with a simple output layer and no fine-tuning.

1 Introduction

1.1 History of fashion plate and its importance

The history of art is an important part of human history. We do not know what drove the first humans to create a work of art and what that work of art was, but we can analyse the vast resources humans left behind to better understand the world and the people in it. Fashion plates are fashion illustrations from ages past,through the exact years of their existence are debated in literature.
Fashion history as a complementary science to history[1], combines the style of clothing with the architecture, painting, sculpture, literature and music of the era. This makes it a rich source of knowledge not only for historians and art historians, but also for philosophers, psychologists and other scientists, as well as for hobbyists, fashion designers and theatre and film costume designers.
Historical sources are an important documentation of human achievements. Iconographic historical sources, along with written artefacts, are the main source about the prevailing fashion of the period. Fashion plates can be used to analyse the clothing of a particular period, as they focus primarily on the depiction of the clothes rather than the people that wore them. These plates depict the prevailing fashions of the period in which they were made, and because they were used as a form of fashion catalog for the industry, they were often a precursor to what became that period's fashion[2]. Modern day examples of this would be fashion magazines and shows presenting the next season's trends. However, the work in the production of these fashion plates was not often attributed to individual artists, nor was the date of creation always recorded. Hence, there is now a need to research the provenance of these fashion plates so that they can provide more useful information about the history of fashion.Dating fashion plates is valuable for its own sake, but also for other forms of historical research. Changes to fashion through the ages helps us to date other historical artifacts such as written sources and paintings.

The digitisation of museums [3], library or other history-related collections is an important goal for the preservation and forward transmission of our heritage. One way to digitise is to use expert knowledge. However, in areas where expert knowledge is very valuable - specialists in fashion history are relatively few in number and their time is therefore limited and valuable - it is worth exploring any means by which the task of digitisation can be automated.

In this case, the idea of using machine learning to digitise collections is a sensible one. While expert knowledge may be needed to verify the results of an automated process, the initial stages can certainly be handled without the costs associated with expert knowledge. Even without definitive verification of the dates of the materials in the fashion plate collection, the information obtained through machine learning could prove valuable - for hobbyists and students, for example - provided the results are reasonably close to the correct date.This task provides many challenges, such as limited data which can be eliviateted with using pre-trained models. Transfer learning is the approach of using pre-trained models of artificial intelligence and adjusting them with additional training for the new problem. In this case this new problem is the recognition of date of origin of fashion plates created in the period from 1820 to 1880.

1.2 In depth analysis of 1860-1862

Dating fashion plates precisely is a difficult task which requires a lot of knowledge about major and minor details of dress. To show how many factors are at play a brief description of the opening years of the 1860s will serve as an illustration[3].

Starting in 1860 the skirts are extremely full, held up on a round crinoline[1] with a few petticoats. They areoften decorated with lots of ruffles. Decorations can be both vertical and horizontal. Some dresses consist of two skirts, one worn over the other, the top one pulled up in a few places to reveal the bottom layer. The sleeves are made in a large pagoda style[2] with white undersleeves showing beneath or consisting of a lot of puffs, the latter appearing on fashion plates exclusively from 1860. The most common bodice hem shape is pointed, followed by princess seams[3]. In day wear there are also some examples of the cutoff between bodice[4] and skirt going straight across. There does not seem to be any limits to how bodices are decorated. Evening bodices sit low off the shoulder with big berthas[5] with sleeves covered by ruffles. In contrast, day bodices end very high.

In 1861 the round crinolines have changed into a slightly more elliptical shape. Ruffles are still very common but unlike the previous year, there are many geometric designs such as Greek keys[6], square crenelations[7], scallops[8]and bands of trimming. The pagoda sleeve shrinks significantly but even in the smaller sleeve the white puff of the undersleeve[9] remains. There is also some variation on day bodice necklines, although the high necklines remain popular there is also the option of lower square necklines filled out by chemisettes[10] or V necklines both filled out and not. Evening bodices remain the same as the year prior.

In 1862 the look becomes more streamlined. There are almost no ruffles, sometimes just a few rows on the bottom (as opposed to a whole skirt covered in them). The most popular decoration is several horizontal bands of trimming going up from the hem and the geometric designs of 1861. The majority of skirt decoration is on the lower third of the skirt with a plain upper portion. Daywear bodice trimmings are mostly absent apart from some decoration on the sleeves with gentle flare(though the white cuff remain). The point of the bodice hem has become a lot gentler or even straight. There is also a new finish to the bodices: two points on either side of the center front opening. Surprisingly there are no princess seam-style bodices despite their popularity in previous years. Evening necklines have become slightly higher with more defined sleeves (a small puff, sometimes covered with ruffles).

1.3 Scope of the research

In this paper I will investigate which model and approach is best for recognizing the date of origin of iconographic sources in the fashion history field. To achieve this I will create my own database of fashion plates and train several different model architectures through the use of four different approaches

[1]crinoline- a round or elliptical understructure for skirts used in the 19th century

[2]Pagoda sleeve- a sleeve fitted at the top and gradually flaring near the hem

[3]Princess seam dress- a dress with seams going through the breast apex to the hem of the dress. That means the dress no longer consists of a separate bodice and skirt

[4]Bodice - the upper part of a dress

[5]Bertha - a decorative element going over the shoulders and across the chest, often covering the top of evening gowns

[6]Greek keys- a decorative border constructed from a continuous line, shaped into a repeated motif

[7]Square crenelations- square teeth or notches of a repeating structure

[8]Scallops- repeated semicircular pattern

[9]Undersleeve- a part of the sleeve attached by the hem that can be easily detached and washed

[10]Chemisette- a piece of fabric which looks like the top of a blouse used to fill out necklines

Figure 1. H. Holbein- The Ambassadors 1533

(described in Experimental Design) and compare their results. However, I will not be creating my own architecture.

1.4 Summary

As discussed above fashion plates are important not only for fashion history but also many other fields of study. That is why it is important to date them accurately. Using transfer learning is one of the possible solutions for speeding up that process. This leads to the question: Which model and approach is best for streamlining recognition of date of origin of iconographic sources in the fashion history field? The methodology used to tackle this query was chosen as it was found to be appropriate for use in small databases, as shown by Zhu et a. [4] and due to the nature of working with surviving historical artifacts, this number is quite limited. Based on my results where the best model achieved a mean absolute error of 2.8 years in a range of 60 years it is a feasible solution.

2 Related work

The study of history is closely related to the history of art. Fashion history is a complementary study to that of mainstream history as the style of clothing stands in close relation to trends in architecture, sculpture, painting, literature, and music. Knowing the history of clothing allows you to accurately estimate the dates of not only the iconography directly related to fashion, but also of any paintings, sculptures and photographs that happen to include such clothing, as fashion is ubiquitous throughout human civilization . The choice of clothing determined by what was available at the time is indicative of the date of origin for many works including scenes of daily life, for example: "The Ambassadors" (Holbein 1533) as shown in Figure 1 can be dated to the Tudor period[5] due to the characteristic sleeves and other elements worn by the two subjects within the piece.

Fashion was and is an integral part of life. It shows current artistic movements, the priorities of the people dressed in that fashion and much more therefore it is important to correctly identify the periods of origin of a given historical source, which allows researchers to place a given artifact in the broader context of history. One way to preserve fashion history is to digitize available collections.

The chronological data on artistic works are often lacking in regard to the names of creator of the artwork,, the style of the clothing and the original date of creation, all of which generates difficulties when it comes to the digitization of existing works. There have been many attempts to address this topic and the topic of art categorization.

There are no publicly available databases of fashion plates from the period 1820-1880, so the database was created using web scraping as further described in Section 4.1 .
The same method was used by Mao et al. [6]to create the database known as Art500k which contains 554,198 files depicting visual arts of various forms. The pieces were gathered from many different sources, including the RijksMuseum, Google art and culture, WikiArt and the Web Gallery of Art and by web scraping 10 different attributes for each image. MySQL formed the basis for the database and the database was published in multiple formats. A two-track triplet network was built, based on the VGG-16 architecture. The team used two different sets of tasks to evaluate: search and annotation; against several solutions with manually created features, as well as variations of VGG-16. They were able to achieve the best results compared to benchmarks for both types of tasks, but found the results of the annotation tasks unsatisfactory and in need of further development.

Another study into the classification of art, Strezoski et al.[7] used more than 2 million images to create a dataset to assign author, classify type, genre and subtype of work, as well as the period of origin of given works. Both classification and prediction were used to determine the period of origin for each work. These techniques were based on information obtained from textual metadata annotations. Performance was significantly affected by different labeling and training schemes used in each of these approaches. In the setting of predictive (regression) years and age of creation were estimated as a function, with an average absolute error loss. For classification, softmax with 84 classes was utilized to assign age of emergence. It was noted that attempting to identify the creation period in centuries is more effective than in years regardless of whether the problem was approached as classification or prediction (regression).

Another study concerning the categorization of art, in this case sculpture, was conducted by El Vaigh et al.[8] in which they described the categorization of the sculpture by selecting the style and period of creation for analysis. 2,665 Buddha statues were used and they compared the results to a benchmark dataset SemArt dataset of 21,382 images with traits such as technique, title, author, date, time frame, type and school.

Using a hybrid approach, they combined Convolutional Neural Networks (CNNs) with knowledge graphs (KGs). They assigned pseudo tags to unlabeled data that was enriched with KGs, to create extended knowledge graphs. An improvement in the classification of SemArt artwork datasets and Buddha statues was noted after using convolutional graph networks.

Among the available articles on the subject of art categorization, the most commonly occurring subject was paintings. Viswanathan[9] was the first to use a CNN to solve this problem and provides one such example of categorizing a set of images with indications of author, genre and style in his 2017 paper, which took 300 images from each of the 57 selected artists as its basis. He also compared 4 approaches present in earlier works: a basic CNN built from scratch, SVM and two iterations of ResNet: trained from scratch, and transfer learning. He achieved the best results with ResNet with transfer learning with a test accuracy for top-3 of 0.898.
Kelek et al.[10]kel provide another such dataset, originating from 17 painters with 46 images each. They used 5 networks to classify the images: Resnet50, Inceptionv3, GoogleNet, Resnet101, and DenseNet, and tested them all with different test coefficients. Inceptionv3 and DenseNet turned out to be the best CNNs, with a slight advantage for DenseNet; however, one noted disadvantage is that the training time for this network is twice as long as for others, and, overall, Inceptionv3 was determined to be the most

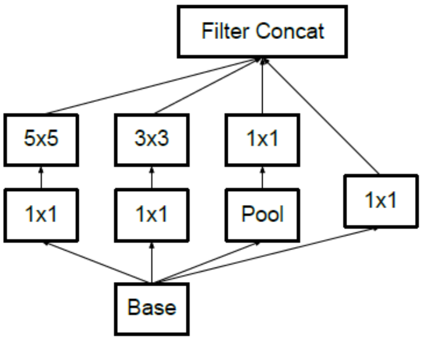

Figure 2. Original Inception module.

efficient of the networks.

There is much research that explores the categorization of art. For the correct digitization of source evidence, it is important to categorize according to the year the work was created. Most studies on art categorization have used CNN. Among the studies I know of, none approaches the problem of art categorization with a focus on historical fashion, a lacuna that the present paper seeks to fill.

3 Background

3.1 InceptionV3

InceptionV3 [11] is a neural network architecture that was introduced by Google. It was designed to enable more complex neural networks while controlling the growth in the number of parameters. It uses fewer than 25 million parameters, a significant difference from its predecessor AlexNet, which uses 60 million parameters. InceptionV3 achieved state-of-the-art results on the ImageNet dataset in 2015, with a top-1 accuracy of 78.95

InceptionV3 is built using convolutional layers, pooling layers, and inception modules. The Inception module in the InceptionV3 network combines multiple parallel convolutional filters of different sizes to capture features at different scales and resolutions. This allows the network to capture both local and global features of the input image while controlling its computational cost.

Inception modules use mini-networks as a building block instead of traditional convolutional filters as shown in Figures 2 and 3. These mini-networks are made up of smaller 1x1 and 3x3 convolutional filters and are used instead of larger convolutional filters, which can be computationally expensive.

Mini-networks use smaller convolutional filters, typically 1x1 and 3x3, instead of larger filters. They have two or more layers, with the first layer reducing the number of input channels and the second layer capturing more complex features. The output is then concatenated and used as the final output. Mini-networks help control computational costs while allowing deep neural networks to capture a wider range of features.

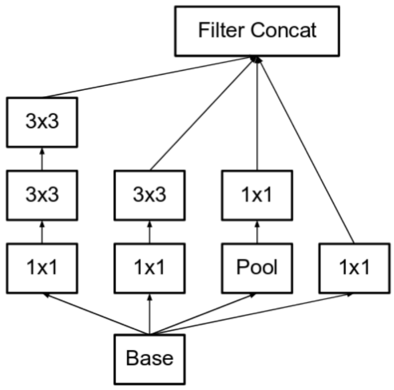

Figure 3. Inception modules where 5×5 is replaced by mini-network

3.2 ResNet101

ResNet101 [12] is a deep convolutional neural network architecture that was introduced by Microsoft Research in 2015. It is part of the ResNet family (ResNet-50, ResNet-101, and ResNet-152). ResNet101 achieved state-of-the-art results on the ImageNet dataset in 2015, with a top-1 accuracy of 78.25

In a residual block, the input to the block is passed through a series of convolutional layers, followed by a shortcut or skip connection that bypasses the convolutional layers and directly connects the input to the output of the block, as shown in Figure 4.

The skip connection (Figure 5) allows the network to learn residual functions that capture the difference between the input and output of the block, rather than directly learning the mapping from input to output. This helps to mitigate the problem of vanishing gradients that can occur in very deep neural networks, which can capture more complex features in images.

ResNet101 contains multiple residual blocks (Figure 6), each consisting of several convolutional layers and a skip connection. By stacking these blocks together, ResNet101 is able to learn complex representations of input images, leading to high accuracy on a variety of computer vision tasks.

3.3 NasNetMobile

NasNetMobile is a neural network architecture that was published by Google in 2018. It is a mobile-friendly architecture, designed to be computationally efficient and suitable for use on mobile devices. NasNetMobile was trained on the ImageNet dataset and achieved a top-1 accuracy of 82.7% [13].

NasNetMobile is built using neural architecture search (NAS), which is a technique for automatically designing neural network architectures. This process can take a long time, but it has the potential to find more efficient and accurate architectures than manually designed processes.

The NasNetMobile architecture has these two main components (Figures 7 8):

Normal cells: These are the main building blocks of the architecture. Each normal cell contains multiple parallel convolutional blocks, each of which has a different dilation rate - the rate of how much the initial matrix is reduced per convolution. This block returns a map of the same dimension. Reduction cells: These are similar to normal cells, but they are used to reduce the spatial resolution of the feature maps. This cell returns a feature map where height and width are halved.

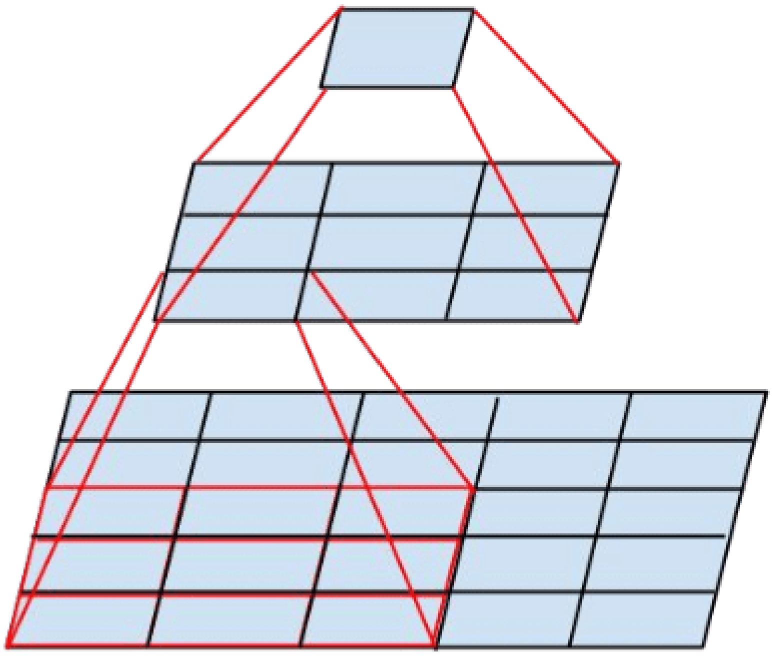

Figure 4. Mini-network replacing the 5×5 convolutions. The lower layer of this network consists of a 3×3 convolution with 9 output units

The convolutional blocks in the normal cells use a combination of 1x1, 3x3, and 5x5 convolutions, as well as max pooling and average pooling. The dilation rates of the convolutional filters are chosen by the neural architecture search algorithm to optimize performance.

3.4 Comparison

ResNet101 and InceptionV3 are both designed using traditional manual architecture design techniques, whereas NasNetMobile uses neural architecture search. The first two both use techniques to enable the training of very deep networks, with ResNet101 using skip connections and InceptionV3 using inception modules and mini-networks. NasNetMobile, on the other hand, is designed to be computationally efficient and suitable for use on mobile devices. As a result, it achieves slightly better performance despite its much smaller size.

Table 1

*Comparison of the models.*The NASNet-Mobile neural network does not consist of a linear sequence of modules.*

Neural Network	Depth[14]	Size [14]	Parameters [14]	Year (Published)	Top-1 acc ImageNet [15]
InceptionV3	48	89MB	24M	2015	78.95%[2]
ResNet101	101	167MB	44.6M	2015	78.25%[2]
NasNetMobile	*	20MB	5.3M	2018	82.7%[2]

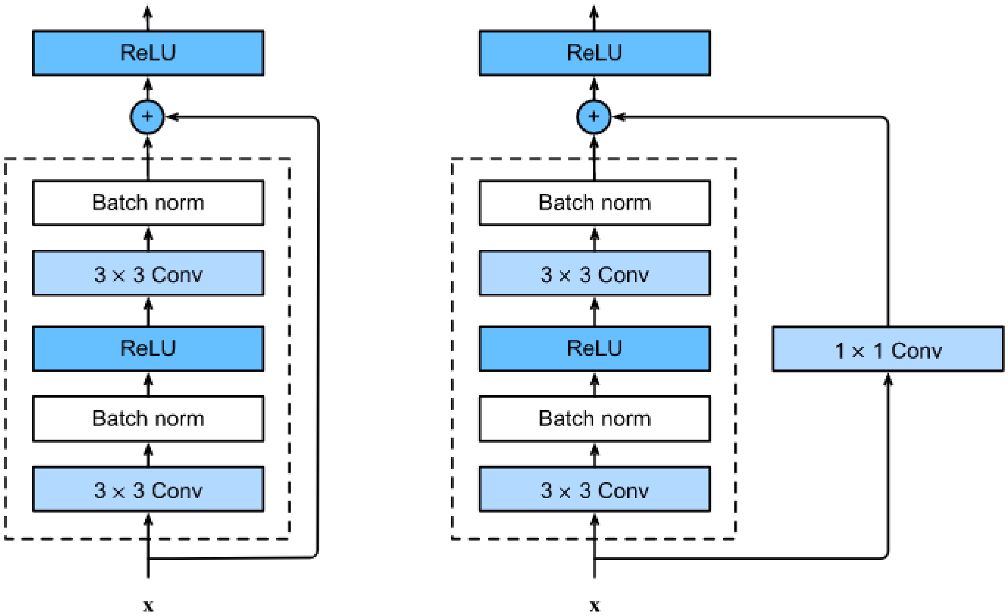

Figure 5. A residual block with skip connection. [11]

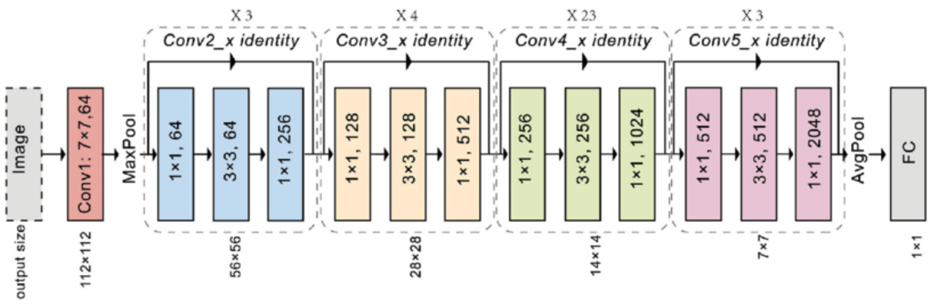

Figure 6. Typical architecture of the 101-layer ResNet. [12]

4 Methodology

4.1 Data

Web scraping was used to obtain the necessary fashion plates since in the initial research for the topic widely available database of fashion plates could not be found. I utilised the Selenium library due to the dynamic nature of the pages. At first I turned to renowned museum websites such as the Metropolitan Museum in New York and Victoria and Albert Museum in London, both highly regarded for their extensive library of fashion history artifacts, along with the University of Washington Library which also has a valuable source of fashion plates that I utilised. The Wikimedia database is an extensive compendium hosted by the non-profit Wikimedia Foundation which houses a wide variety of fashion plates, amongst other resources. I also turned to an online community on Etsy to download additional fashion plates. There were a number of other sites that provided a comparatively small sample size of fashion plates, with only a few dozen images per page. While I was gathering the data I was planning to use categorization rounded to the decade each fashion plate is from. That means that when the date was uncertain I

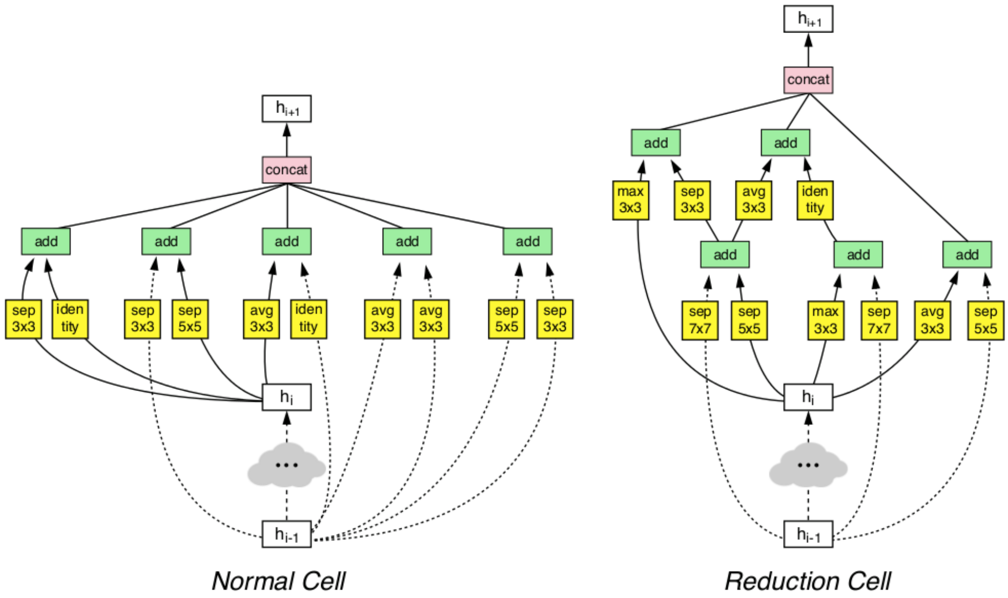

Figure 7. The previous activation's hidden state is passed as an input (white cells), and the resulting branches' concatenation creates the output (pink).

rounded it to the earlier date which caused some mislabeling in the data. For future research, it would be advisable to take the average of the earliest and latest possible dates provided in order to obtain more accurate information. Afterward, an exact timeframe was chosen based on the number of samples available. All the gathered data were manually combed through to discard pictures that did not focus on dresses (such as male fashion, hats, and other accessories) as well as outfits of regional dress or specialty uniforms. After that process 18 807 fashion plates were approved and 2 525 rejected. The exact breakdown of the final dataset of fashion plates can be found in Table 1 and Figure 9.

Finally, the dataset was divided into training(9937 fashion plates), validation(1242 fashion plates) and

Source	**number of fashion plates**
etsy.com	6595
other	53
metmuseum.org	2002
vam.ac.uk	1500
lib.washigton.edu	157
wikimedia.org	2114
SUM	**12421**

Table 2

The place of origin of web scraped fashion plates

test(1242 fashion plates) sets randomly.

4.2 Method

To prepare the pictures all were resized using the Tenserflow (Figure 10) library and then the data were augmented, though only in the training set. The augmentation consisted of randomly flipping (horizontally and vertically),and changing the rotation as well as the contrast.

I used the methods presented by Zhu et al. [3]. Having a database of 200 images focused on the qual-

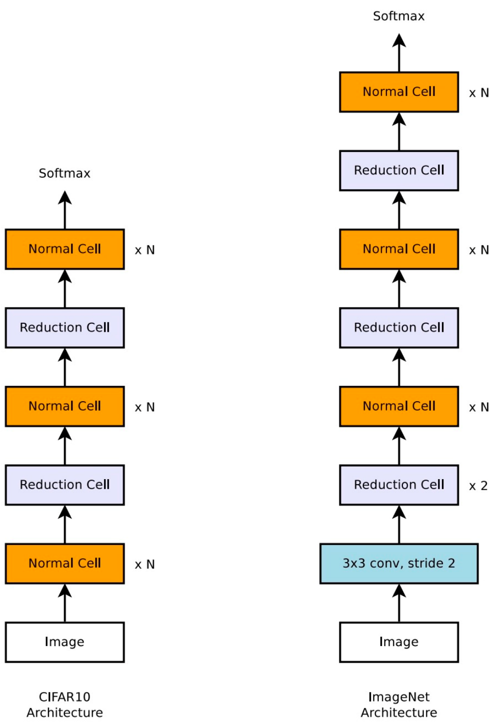

Figure 8. NasNetMobile architecture has Normal and Reduction Cells that repeat. Both in its CIFAR and ImageNet implementations.

ity of plastic pellets, they analyzed various model setup methods to choose the most optimal approach. They used ResNet-152 and ResNet-18 to classify the limited data resources.They used four different approaches

Methods used:

- Fine-tuning One additional layer
- Transfer learning One additional layer
- Fine-tuning A Dense classifier with a single unit
- Transfer learning A Dense classifier with a single unit

"One additional layer" can be understood as the addition of an extra layer before the decision-making neuron (number of neurons determined by the BayesianOptimization function)
"A dense classifier with a single unit" means adding a single neuron with an activation function.
Fine-tuning requires taking pre-trained models with small additions and training them in their entirety but with a small learning rate to avoid destroying the knowledge embedded in the network.
Transfer learning takes those pre-trained models and attaches only a small section to be trained and all other weights are frozen.
In addition to this I chose the best model setup from the ones trained and trained a similar model but this time on categorization, not regression. For that I chose the best combination out of the regression approaches.

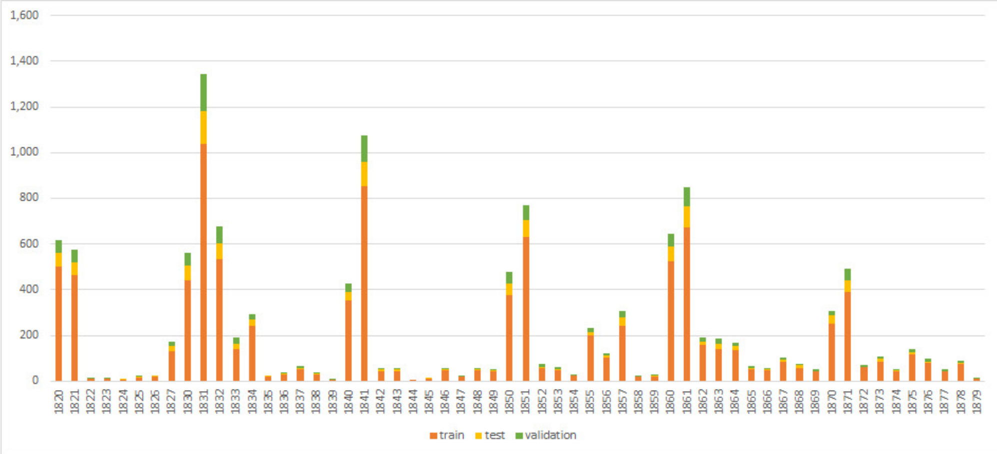

Figure 9. Distribution of gathered fashion plates over the analyzed period

Figure 10. The difference in quality of a few fashion plates before and after resizing

	Fine-tuning	Transfer learning
One additional layer		
A dense classifier with a single unit		

5 Experimental Design

The loss function used for measuring the performance of all the models is the mean squared error and is the metric monitored to help gauge the performance of the network during training. Another metric monitored is the mean absolute error. The adam optimizer was used. Batch size was set to 16.
The early stopping point is utilized to stop the neural network from overfitting on train data which causes a drop in performance of any data outside of the dataset, and was reached in all experiments and was based on validation loss with patience = 4.
To apply regression to all models I connected a dense layer with a sigmoid activation function. Classification uses each year as its own class whilst regression uses a normalised (0,1) value of all the years.

6 Results and Analysis

6.1 General Results

Presented below are the results of the study with the averages calculated over all of the approaches used for each model.

The best results for the Resnet101 model came from the configuration of using an extra layer and using fine tuning which resulted in a Mean absolute error (years) of 4.79 and Accurate predictions accounting for 9.26% of all predictions. The average scores of ResNet101 were Mean absolute error in years of 6.3325 and accurate predictions of 6.14%.

For the NASNetMobile model, the best result was through the absence of an additional layer and the use of fine tuning, resulting in a Mean absolute error (years) of 2.8 and Accurate predictions of 18.68%.The average scores of NASNetMobile were Mean absolute error in years of 6.8 and accurate predictions of 9.14%.

The use of the InceptionV3 model in the juxtaposition of the use of additional layer and transfer learning achieved 3.42 Mean absolute error (years) accuracy and 13.85% Accurate predictions. The average scores of InceptionV3 (without the rerun) were Mean absolute error in years of 19.14 and accurate predictions of 6.06%.
The highest average results were achieved by NASNetMobile followed by ResNet101 and InceptionV3. What stands out the most are two outliers of InceptionV3. It is unclear why this occurred but one can speculate it has something to do with random factors such as data augmentation but even retraining one of those outlier combinations did not yield much better results.

Despite prior concerns due to the uneven nature of the data, the best model seemed to learn how to recognise most of the years, not only the ones with high resources.

To further analyze the results I created an algorithm that covers consecutive parts of the image and runs inference for each of the resulting images. That allows me to see how important each square is for decision-making by calculating the difference between the actual year and the year guessed with part of the image covered. To better illustrate the results I created a heatmap as shown in Figure 10. From the saliency map one can notice that the majority of the focus is on the correct object, i.e. the dresses, but some of the focus is concentrated on the top of the image where there is no valuable data. Although the network mostly focuses on the correct object, the specific areas of the dresses it looks at are clearly different than what a human with extensive knowledge of the subject would concentrate on.

6.2 Qualitative Analysis of the worst prediction

To take a closer look at the results of the best model I analyzed the fashion plate with the worst prediction, shown in Figure 12. As one can see the silhouettes are typical for the "natural form era" (starting from around 1875) with slim skirts often ending with a train, the train being a term for the part of a skirt trailing behind the individual wearing it. This style of dress has very few examples available compared to other styles present in the dataset as the existence of this style of dress begins within the last four years of the dataset from the sixty year period analysed. For example, while the Crinoline era (1856-1869) is captured in its entirety, as the beginning and end of this era are contained within the dataset, whereas the Natural Form era is not, ending in 1882. The year of origin of this fashion plate is 1879 yet the model predicted

Model	additional layer	approach	Mean absolute error (years)	Accurate predictions	Standard deviation
ResNet101	0	fine tuning	6.75	5.395%	9.08
ResNet101	0	transfer learning	6.53	4.992%	8.70
ResNet101	1	fine tuning	4.79	9.259%	6.80
ResNet101	1	transfer learning	7.26	4.911%	9.74
NASNetMobile	0	fine tuning	2.80	18.680%	4.86
NASNetMobile	0	transfer learning	10.20	2.496%	12.94
NASNetMobile	1	fine tuning	3.60	13.040%	5.70
NASNetMobile	1	transfer learning	10.60	2.335%	13.44
InceptionV3	0	fine tuning	4.34	10.390%	6.49
InceptionV3	0	transfer learning	34.40	0.000%	16.31
InceptionV3	1	fine tuning	34.40	0.000%	16.31
InceptionV3-second attempt	1	fine tuning	25.60	4.911%	16.48
InceptionV3	1	transfer learning	3.42	13.850%	5.67
NASNetMobile classification	0	fine tuning	25.60	4.670%	17.84

Table 3

The results of all the training conducted

1849. There are a number of potential causes, one such reason being that it might have read it wrong due to the fact that the lady on the right is putting on a coat of the same color as her dress, which gives the illusion of a wider skirt, especially in lower resolutions. To further test this theory I manually masked out the coat, shown in Figure 14, and repeated the inference getting the year 1856 which is more accurate than the previous prediction but is still out by 23 years. This suggests a further issue: that 1879 is the end of the analysed period which limits the samples available around that particular year. The saliency map, in Figure 15, of this image shows the correct areas of focus which indicates that the algorithm learned to focus on the important parts of the image. The second lady is wearing a coat that obstructs the typical adornment of the era, such as geometric shapes created by ribbons and other decorations like lace. Lastly, the fashion plate has a lot of background elements, unlike some other fashion plates which show only one person on white background. A dress typical of 1849 can be seen in Figure 13.

7 Discussion

Achieving the results of mean absolute error of 2.8 years in a range of 60 years shows potential for using the results of this study in practice. Such a narrow window of error can help speed up the process of identifying the date of origin of fashion plates and can help amateurs research the correct periods suited to their specific areas of interest. Furthermore, since museums delegate the identification of artifacts to experts, with an estimate of the date of origin they can spend less time choosing the most appropriate expert; additionally, once it is in the hands of an expert 2.8 years of error can be useful in choosing the right time to further analyze. Finally this AI can be utilized as a training tool for those who wish to join this field of study.

To better understand why some approaches failed it might be beneficial to monitor some random values, such as which fashion plates were flipped, to better understand the differences between training in future work. One way of executing this would be to save all those parameters and names into a file. Though it is not standard practice to monitor this information it could provide insight as to why a certain model or

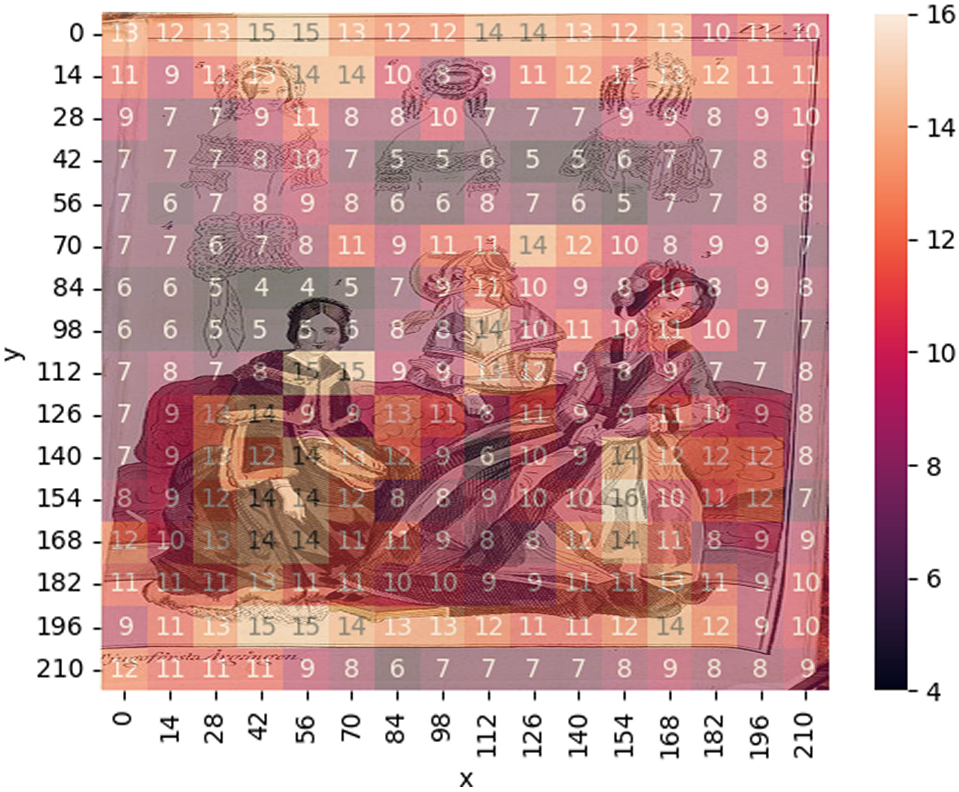

Figure 11. The attention map(the higher the difference/lighter the color, the more important the square) generated from the best model on one of the fashion plates with correct prediction

approach failed to achieve better results.
Another idea for future work would be training models in each approach multiple times to achieve better estimates of their capabilities.
Additionally, it is worth conducting a similar study on the more difficult subject of surviving fashion artifacts such as bodices, skirts, and full ensembles. This task is more challenging for many reasons, such as the limited number of artifacts, and that the photographs of these artifacts are not only of varying quality but also taken in different conditions, all of which can greatly impact the training.

Other factors worth considering are augmenting the colour and saturation of the images. Also due to the time constraints of this study one batch size was chosen for all models which may have negatively impacted some of them - changes in the batch size are worth investigating. Another approach that is popular when it comes to working with art artifacts in machine learning is the use of knowledge graphs, which provide expert knowledge into the system which could greatly improve the reliability of the results.

Lastly, this study limited the origin of fashion plates to between the narrow timeframe of 1820-1880. Future studies can expand said period to cover even more artifacts. One of the challenges of expanding the range of study would be the sparsity of available data - the accuracy of determined dates also diminishes with the regression throughout the preserved historical record.

8 Conclusion

Returning to the research question posed at the beginning, "Which model and approach is best for streamlining recognition of date of origin of iconographic sources in the fashion history field?", we can say that NASNetMobile provided the best results, achieving a mean absolute error of 2.8 years in a range of 60 years. The best approach utilised fine tuning without an additional layer, as opposed to the second best result which utilized fine tuning with an additional layer for a mean absolute error of 3.6 years.

References

[1] A. Sieradzka, Kostiumologia Polska jako nauka pomocnicza historii. 2013.

[2] C. Ness, 'April Calahan, Fashion Plates: 150 Years of Style, ed. by Karen Trivette Cannell', Costume, vol. 51, no. 1, pp. 135–136, Mar. 2017, doi: 10.3366/cost.2017.0011.

[3] All About 1860's Fashion // What did Civil War-era fashion look like?, (Nov. 12, 2022). Accessed: May 08, 2023. [Online Video]. Available: https://www.youtube.com/watch?v=0PLu18-Y0rM

[4] N. Mikhaila and J. Malcolm-Davies, The Tudor Tailor: Reconstructing Sixteenth-Century Dress. Batsford, 2006. p.13

[5] H. Mao, M. Cheung, and J. She, 'DeepArt: Learning Joint Representations of Visual Arts', in Proceedings of the 25th ACM international conference on Multimedia, Mountain View California USA: ACM, Oct. 2017, pp. 1183–1191. doi: 10.1145/3123266.3123405.

[6] G. Strezoski and M. Worring, 'OmniArt: A Large-scale Artistic Benchmark', ACM Trans. Multimed. Comput. Commun. Appl., vol. 14, no. 4, p. 88:1-88:21, Październik 2018, doi: 10.1145/3273022.

[7] C. B. El Vaigh, N. Garcia, B. Renoust, C. Chu, Y. Nakashima, and H. Nagahara, 'GCNBoost: Artwork Classification by Label Propagation through a Knowledge Graph', in Proceedings of the 2021 International Conference on Multimedia Retrieval, in ICMR '21. New York, NY, USA: Association for Computing Machinery, Sierpie 2021, pp. 92–100. doi: 10.1145/3460426.3463636.

[8] N. Viswanathan, 'Artist Identification with Convolutional Neural Networks', p. 8.

[9] C. Szegedy, V. Vanhoucke, S. Ioffe, J. Shlens, and Z. Wojna, 'Rethinking the Inception Architecture for Computer Vision'. arXiv, Dec. 11, 2015. doi: 10.48550/arXiv.1512.00567.

[10] K. He, X. Zhang, S. Ren, and J. Sun, 'Deep Residual Learning for Image Recognition'. arXiv, Dec. 10, 2015. Accessed: Mar. 26, 2023. [Online]. Available: http://arxiv.org/abs/1512.03385

[11] '8.6. Residual Networks (ResNet) and ResNeXt — Dive into Deep Learning 1.0.0-beta0 documentation'. https://d2l.ai/chapter_convolutional-modern/resnet.html (accessed May 08, 2023).

[12] Y. Tong, W. Lu, Q. Deng, C. Chen, and Y. Shen, 'Automated Identification of retinopathy of prematurity by image-based deep learning', Eye Vis., vol. 7, p. 40, Aug. 2020, doi: 10.1186/s40662-020-00206-2.

[13] B. Zoph, V. Vasudevan, J. Shlens, and Q. V. Le, 'Learning Transferable Architectures for Scalable Image Recognition'. arXiv, Apr. 11, 2018. doi: 10.48550/arXiv.1707.07012.

[14] 'Pretrained Deep Neural Networks - MATLAB Simulink'. https://www.mathworks.com/help/deeplearning/ug/pretrained-convolutional-neural-networks.html (accessed May 08, 2023).

[15] 'Papers with Code - ImageNet Benchmark (Image Classification)'. https://paperswithcode.com/sota/image-classification-on-imagenet?metric=Top%5C%205%5C%20Accuracy (accessed May 08, 2023).

[16] W. Zhu, B. Braun, L. H. Chiang, and J. A. Romagnoli, 'Investigation of transfer learning for image classification and impact on training sample size', Chemom. Intell. Lab. Syst., vol. 211, p. 104269, Apr. 2021, doi: 10.1016/j.chemolab.2021.104269.

Figure 12. The fashion plate with worse prediction in the best model, actually from 1879, predicted as 1849

Figure 13. Masked fashion plate with the worst prediction

Figure 14. A fashion plate from 1849

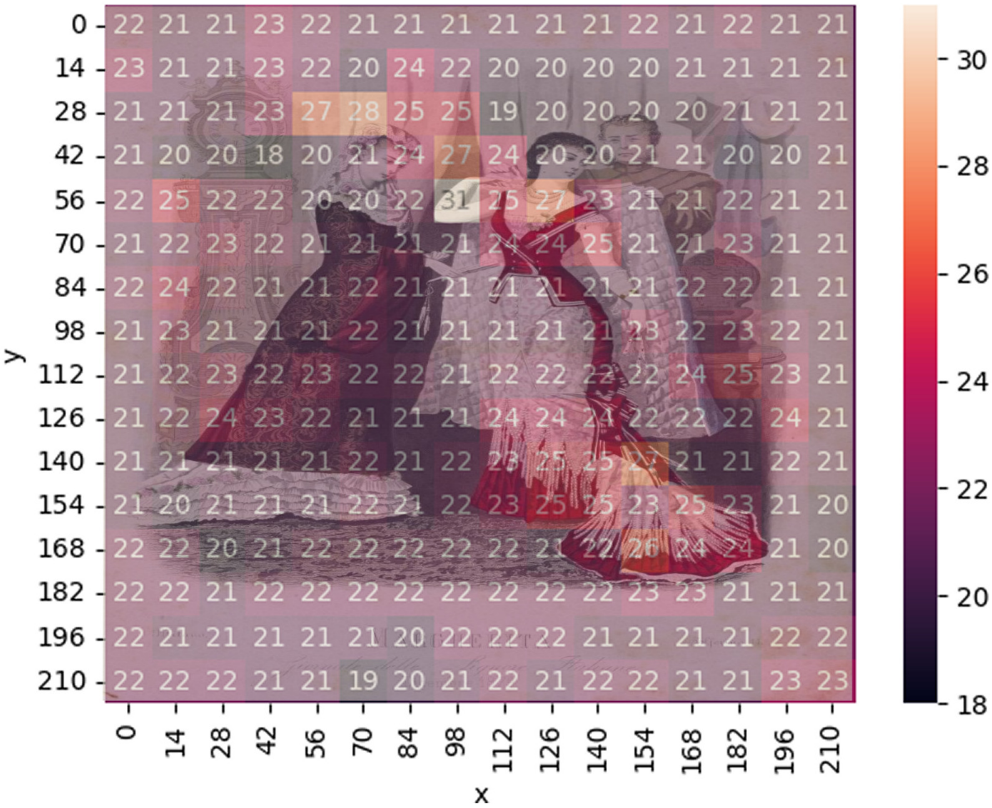

Figure 15. Saliency map for the worst predicted fashion plate