

Article

Automatic Gender and Age Classification from Offline Handwriting with Bilinear ResNet

Irina Rabaev ^{*,†} , Izadeen Alkoran, Odai Wattad [†]  and Marina Litvak ^{*,†} 

Software Engineering Department, Shamoon College of Engineering, 56 Bialik St., Be'er Sheva 8410802, Israel

* Correspondence: irinar@ac.sce.ac.il (I.R.); marinal@ac.sce.ac.il (M.L.)

† These authors contributed equally to this work.

Abstract: This work focuses on automatic gender and age prediction tasks from handwritten documents. This problem is of interest in a variety of fields, such as historical document analysis and forensic investigations. The challenge for automatic gender and age classification can be demonstrated by the relatively low performances of the existing methods. In addition, despite the success of CNN for gender classification, deep neural networks were never applied for age classification. The published works in this area mostly concentrate on English and Arabic languages. In addition to Arabic and English, this work also considers Hebrew, which was much less studied. Following the success of bilinear Convolutional Neural Network (B-CNN) for fine-grained classification, we propose a novel implementation of a B-CNN with ResNet blocks. To our knowledge, this is the first time the bilinear CNN is applied for writer demographics classification. In particular, this is the first attempt to apply a deep neural network for the *age* classification. We perform experiments on documents from three benchmark datasets written in three different languages and provide a thorough comparison with the results reported in the literature. B-ResNet was top-ranked in all tasks. In particular, B-ResNet outperformed other models on KHATT and QUWI datasets on gender classification.

Keywords: automatic handwriting analysis; gender classification; age classification; writer's demographics classification; bilinear CNN; bilinear ResNet



Citation: Rabaev, I.; Alkoran, I.; Wattad, O.; Litvak, M. Automatic Gender and Age Classification from Offline Handwriting with Bilinear ResNet. *Sensors* **2022**, *22*, 9650. <https://doi.org/10.3390/s22249650>

Academic Editor: Jungpil Shin

Received: 19 October 2022

Accepted: 5 December 2022

Published: 9 December 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

It was previously shown that handwriting varies according to several factors, including demographic and geographical region [1]. Handwriting gender and age classifications are of great interest in various fields, including psychology, historical document analysis, and handwriting biometrics. Psychological investigations have confirmed that gender [2,3] and age [4,5] can be classified based on a number of features in handwriting. Typically, male handwriting is more angular, disorderly, and slanted than female handwriting, which is more regular, ordered, and round. The authors of [5] showed a significant effect of age on writing performance. The research of Marzinotto et al. [6] reported the correlation between age and handwriting styles (length of strokes, pressure, stability between words, number of pen-ups, etc.).

This paper focuses on two tasks—automatic gender and age classification from handwriting. Although there exist various automatic approaches for gender and age prediction from handwriting, the challenges of these tasks are illustrated by the comparatively low performances of the proposed models. In addition, despite the success of CNN in classifying gender, deep neural networks were never used for the purpose of classifying *age* from handwriting. Additionally, the majority of the publications that have been published in this field concentrate on the English and Arabic languages. In this work, in addition to Arabic and English, we perform experiments on documents written in Hebrew, which has received a lot less attention.

Based on the success of the bilinear convolutional neural network (B-CNN) for fine-grained classification, we offer a novel implementation of a B-CNN [7] using ResNet instead of VGG blocks. In our previous study [8], we compared multiple deep CNNs on the gender classification task, where VGG demonstrated poor performance. Therefore, in this work, we decided to replace VGG blocks with ResNet, which showed better performance and was relatively easy to integrate within B-CNN. To the best of our knowledge, this is the first time that the B-CNN has been utilized for the classification of writer demographics. In particular, though deep learning methods were previously successfully applied for gender classification from handwriting, this is the very first time that a deep neural network is employed for age classification.

The main contributions of our work are summarized as follows: (1) We modify and apply B-CNN for document image classification; (2) This is the first attempt to apply deep learning models for *age* classification from handwritten document images; (3) We substitute VGG blocks in the B-CNN architecture by ResNet due to the VGG poor performance in our previous study on gender prediction from handwriting [8]; (4) We perform thorough experiments on three benchmark datasets consisting of handwritten documents in three different languages: Arabic, English, and Hebrew. We show that when enough training data are available, B-ResNet outperforms other methods. Namely, on the gender classification task, B-ResNet outperformed other models on the KHATT and QUWI datasets.

The rest of the paper is organized as follows. Section 2 presents the related work and overviews the B-CNN architecture. In Section 3, we describe the proposed methodology, the datasets used for the study, and the experimental settings. Section 4 presents the comparative results of the study. Finally, Section 5 summarizes and concludes the proposed research and provides future directions.

2. Background and Related Work

In this section, we review the related work in the field of automatic handwriting processing. Additionally, we overview the B-CNN model and explain the motivation behind choosing it for the handwriting analysis task.

2.1. Gender Classification

Early works on gender classification applied unsupervised methods [9,10], using Adaptive Multi-Gradient (AMG) [9], and Multi-Gradient Directional (MGD) [10] features.

Since 2012, traditional machine learning methods in general and Support Vector Machines (SVMs) in particular [11–23], have become most popular. Except for SVMs, such models as Decision Trees and their ensembles (Random Forests or AdaBoost) [11,13,21,24], shallow Artificial Neural Networks [11,12,22,25], Regressions [20], Naïve Bayes [21], K-nearest neighbors [11,26], Fuzzy Rule-Based Classification [16], and Discriminant Analysis [21] were applied. We also observed that attention had been paid lately to ensemble approaches [20], where several different classifiers are combined to create a master model. The majority of the aforementioned models were applied upon textural [9,11–13,15–18,25] and a combination of textural and shape features [14,22,23,27–31]. The best accuracy rates—between 77% and 82%—were achieved by the SVM classifiers with textural features [12,16,17,27].

Deep models based on Convolutional Neural Networks started to appear in gender classification works around 2018. Deep neural networks were applied as feature extractors [21], and also end-to-end pipelines, including both feature selection and classification layers [8,32,33]. The main advantage of deep networks is their ability to learn features automatically without manual engineering. In addition, CNNs have been shown to be on par or even outperforming other classifiers on gender classification task [8,33,34]. Due to their benefits in terms of performance and usability, deep networks have recently emerged as a leader in various computer vision applications, including handwriting analysis.

2.2. Age Classification

In contrast to the gender classification task, not many works reported on automatic age classification, while in most of them age was only one of many demographic features identified from handwriting documents.

Bouadjeneq et al. [15] applied an SVM classifier on two gradient features for a writer's gender, handedness, and age range prediction. Three SVM predictors, each applied on a specific data feature, were subsequently combined in [16,35] to identify a writer's gender, age range, and handedness. Emran et al. [36] investigated different classifiers—K-Nearest Neighbors, Random Forests, and SVM—using various visual appearance features for the prediction of a writer's age, gender, and handedness.

Only a few works developed models solely for age prediction. Upadhyay and Singh [37] studied the estimation of age through handwriting characteristics in females and found that such characteristics as slant, alignment, spacing, hesitation marks, tremor, and speed are really valuable and helpful for age determination. Zouaoui et al. [38] investigated the co-training approach for age range prediction from handwriting analysis. The authors proposed several descriptors for feature generation and applied an SVM predictor for classification. Basavaraja et al. [39] proposed a new unsupervised method for age estimation using handwriting analysis with Hu invariant moments, disconnectedness features, and k-means clustering.

In [40], the efficacy of using the dynamic features generated by users of smartphones and tablets to automatically identify their age group was examined. The study with the KNN classifier provides evidence that it is possible to detect user age groups based on the words they write with their fingers on touchscreens. Research in [41] applied SVM and Random Forests to automatically classify people as adults or children based on their handwritten data, collected using a pen tablet.

The best accuracy (up to 81%) was achieved by the SVM classifier with textural features [16], leaving much room for performance improvement in age prediction from handwriting. As can be seen, all works utilized feature engineering in conjunction with conventional classifiers. Deep learning algorithms for age classification have not been used in any research.

2.3. Bilinear CNNs

Bilinear Convolutional Neural Networks (B-CNNs) are a straightforward and efficient architecture for fine-grained visual recognition that, according to Lin et al. [7], generalize a number of orderless texture descriptors, including the Fisher vector and VLAD. These networks capture localized feature interactions in a translationally invariant manner by representing an image as a pooled outer product of features derived from two CNNs. The outer product is produced by the bilinear pooling layer. The idea behind this layer is that the calculated feature interactions enable us to recognize more precise image details. By utilizing the higher-order information acquired in the form of pairwise correlations between features, bilinear pooling offers an efficient method to fuse representative features [42]. B-CNNs are end-to-end trainable and fall within the category of orderless texture representations.

Because textural features were widely used in handwriting analysis [9,11–13,15–18,25], we hypothesized that B-CNN would improve the performance of gender and age prediction from the handwritten text.

The study in [7] compared B-CNNs to both exact and approximate variants of deep texture representations and explored the accuracy and memory trade-offs they offer. As CNN blocks, authors used VGG networks. The key finding was that different outer-product representations perform exceptionally well on a variety of fine-trained, texture, and scene identification tasks. A visualization of B-CNNs in [7] revealed that these models accurately represent objects as texture and that their units are associated with localized features helpful for fine-grained recognition.

3. Methodology

With the advancement of storage and GPU capabilities, approaches based on deep learning have become increasingly popular. Neural networks are capable of learning very powerful and perhaps counterintuitive features that human specialists may overlook. For this reason, we chose to adopt a deep neural network developed for fine-grained classification. We begin with a description of our innovative B-ResNet implementation, followed by a detailed description of the datasets and experimental settings.

3.1. B-ResNet

Handwriting classification is challenging due to small inter-class and large intra-class variations. Handwriting of different persons might share similar visual characteristics, while documents written by the same person can have different appearances due to various external and behavioral factors. Therefore, the proper extraction of discriminating local features describing handwriting is crucial.

In this work, we adopt Bilinear Convolutional Neural Networks (B-CNNs) [7] for fine-grained feature extraction and further classification. Inspired by insights derived from an extensive evaluation of B-CNNs on multiple tasks and datasets in [7], we decided to experiment with B-CNN architecture and apply it to the gender and age classification tasks.

Figure 1 shows the architecture of B-CNN, where two parallel CNN blocks generate representation vectors for an input picture that are later concatenated to one bilinear vector, which is fed to the classification layers.

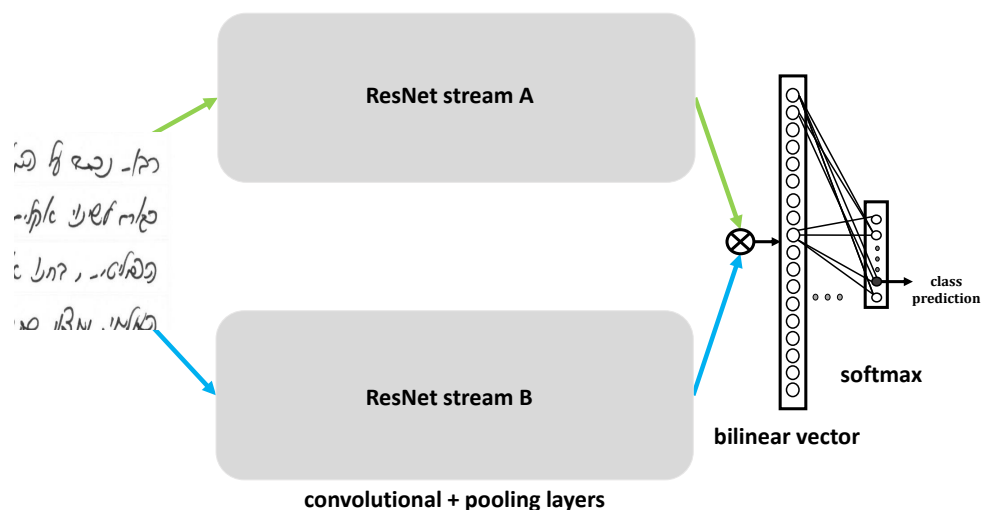


Figure 1. B-ResNet architecture. The input image patch contains Hebrew handwriting.

We first experimented with the original implementation of B-CNN that uses two VGG [43] parallel blocks. However, due to the poor results, we decided to replace VGG with another model. Based on our previous experience [8] with deep networks, where ResNet [44] outperformed VGG, we substituted VGG in two branches by ResNet. The well-known advantage of ResNet is the skip connections that address the vanishing gradient problem. Moreover, ResNet was relatively easy to integrate within B-CNN as compared to other deep-learning models. As expected, the combination of two identical ResNet networks truncated at the last fully connected layer outperformed the combination of VGGs. Therefore, we report its scores in Section 4. We denote our model by B-ResNet for the rest of the paper.

B-ResNet captures translationally invariant localized feature interactions by representing an image as the outer product of features extracted from two parallel ResNets. The bilinear pooling layer produces the outer product. We perform signed square-root normalization, followed by L2 normalization, which, according to [7], improves performance. For classification, we use the softmax function. B-ResNet is trainable end-to-end and naturally

suitable for orderless texture representation. Two ResNets share all parameters; therefore, they have the same memory overhead and runtime as a single network. Two ResNet modules share no computations and are pretrained on ImageNet.

To the best of our knowledge, our work is the first attempt to apply the bilinear CNN model to a handwriting analysis with the purpose of gender and age classification. Moreover, this is the first work on age identification using deep learning models.

3.2. Datasets

For the experiments, we used KHATT, QUWI, and HHD datasets, which are publicly available. Together these datasets consist of documents written in three different languages: Arabic, English, and Hebrew. The KHATT dataset provides gender and age labels; the QUWI and the HHD datasets provide gender labels. Below we describe each one of the datasets, and Figure 2 illustrates sample images from each.

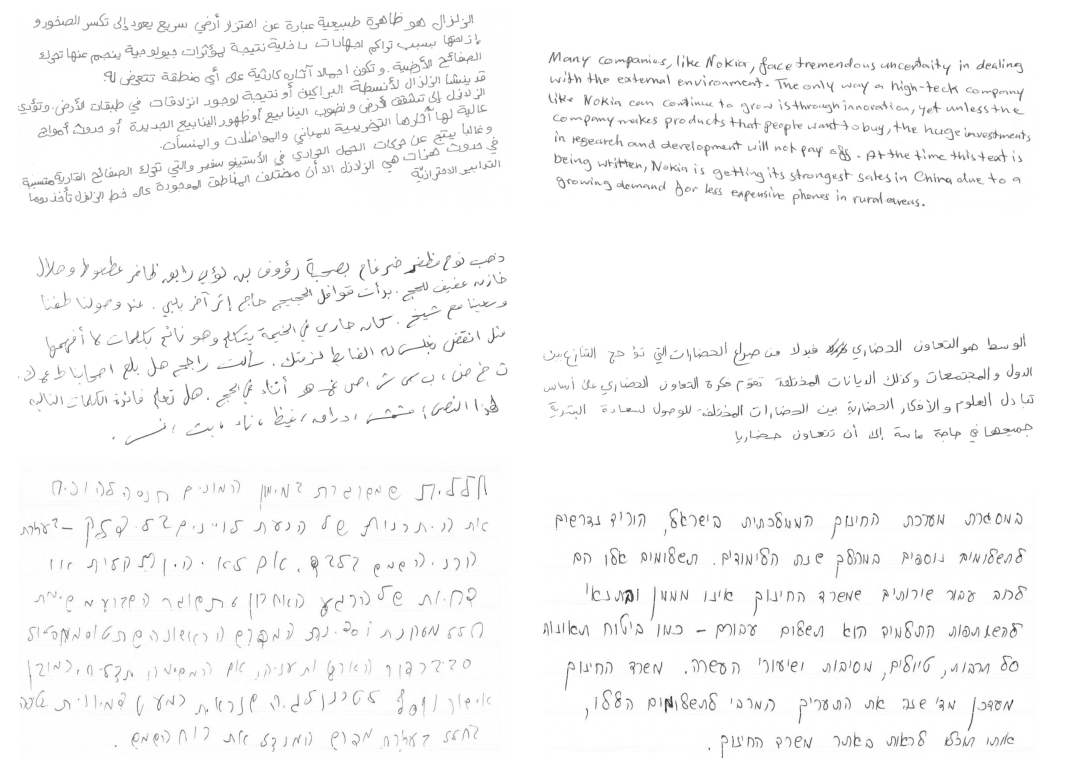


Figure 2. Sample images from the QUWI dataset (top) with Arabic and English handwriting, the KHATT dataset (middle) with Arabic handwriting, and the HHD dataset (bottom) with Hebrew handwriting.

The KHATT (KFUPM Handwritten Arabic Text) dataset (<http://khatt.ideas2serve.net/KHATTDownload.php>, accessed on 1st March 2022) [45,46] consists of handwritten forms written in Arabic by 1000 writers. The ground truth contains gender (male/female) and age group (“<15”, “16–25”, “26–50”, “>50”). The forms were filled in mostly by high school and university students. Each writer contributed five paragraphs: two randomly selected paragraphs from 12 categories, two minimal text paragraphs covering all Arabic letter forms, and one free-text paragraph.

The QUWI (Qatar University Writer Identification) dataset was introduced in [47]. It contains handwritten documents in Arabic and English. The full dataset contains the handwriting of 1017 writers; each writer contributed four pages—two written in English and two in Arabic. The subsets of the QUWI dataset were used in several competitions—ICDAR 2013, ICDAR 2015, and ICFHR 2016 [27–29]. In this work, we utilized the ICDAR2013 subset of QUWI, which is publicly available on Kaggle (<https://www.kaggle.com/c/icdar2013-gender-prediction-from-handwriting/data>, accessed on

21 January 2021). The ICDAR 2013 subset of QUWI includes documents written by 475 writers—221 males and 254 females—and is divided into training and test sets with 282 and 193 writers respectively. There are three alternative configurations used: training and testing on English samples, training and testing on Arabic samples, and training and testing on samples written in both English and Arabic.

The HHD (Handwritten Hebrew Dataset) was introduced in [48]. The HHD_gender (<https://doi.org/10.5281/zenodo.4729908>, accessed on 30 April 2021) subset of the HHD dataset includes 702 forms, filled in Hebrew by 351 men and 351 women. Each participant volunteered demographic information and copied a paragraph of text that was printed above the text field. The form was selected at random from 50 variations and contained an average of 62 words. The HHD_gender dataset is split into training (80%), validation (10%), and test (10%) sets.

3.3. Experimental Settings

In our initial experiments, we utilized the bilinear CNN provided with the [7] paper (<http://vis-www.cs.umass.edu/bcnn>, accessed on 18 January 2022). However, the classification accuracies for both tasks were low (around 50%). We hypothesized that the VGG model used in each branch of the bilinear CNN is not the ideal option for our tasks. The original model was built and trained for general image fine-grained classification, e.g., to differentiate between different bird species or car models. In our domain, we work with document images that are very different from general images. In addition, in our previous study [8], VGG demonstrated poor performance on the gender classification task in comparison to several deep learning models. We decided to replace the VGG model in each branch of B-CNN with ResNet, which showed much better performance and was relatively easy to integrate within B-CNN.

The models were trained on patches extracted from document images. For accurate classification, a document image patch needs to include enough handwriting features. Based on our earlier work [8], a patch with three to four text lines should be sufficient. In our experiments, a patch size of 400 × 400 was sufficient. The patches were extracted by moving a sliding window at a stride of 200 pixels in vertical and horizontal directions. The average number of extracted patches was between 16 and 19. The patches were resized to the input size 448 as expected by the ResNet implementation and normalized with respect to mean and standard deviation. The weights of the two ResNet blocks were pretrained on ImageNet, while the last added layers were randomly initialized. We used a learning rate of 0.1, weight decay 1×10^{-8} , and SGD optimizer. Classification results were evaluated by page-level accuracy, which is computed by taking the majority vote on the predictions of all patches from the page. Figure 3 illustrates the classification pipeline.

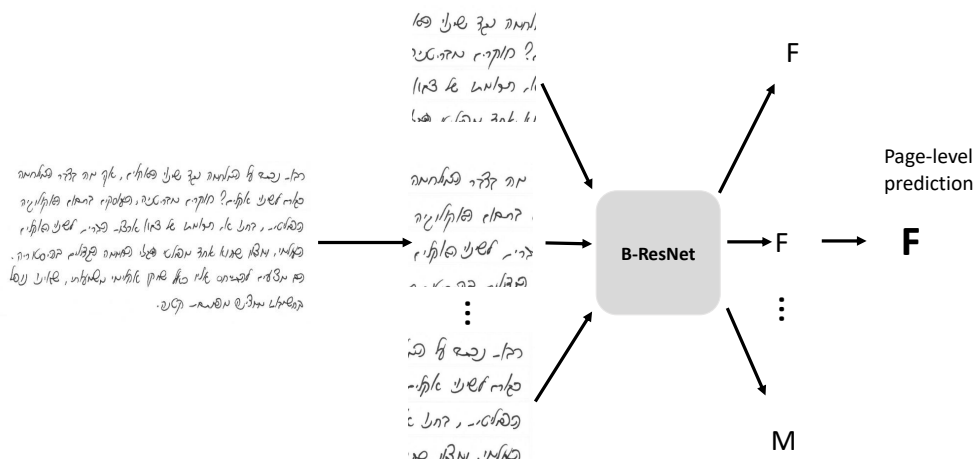


Figure 3. The pipeline of the classification procedure. F and M stand for female and male, respectively. The input image contains Hebrew handwriting.

4. Results and Discussions

For gender classification, we experimented with the KHATT, QUWI (ICDAR 2013), and HHD datasets. For the age classification, we utilized the KHATT dataset. To make a consistent comparison with the results published in the literature, in all the experiments, we followed the official split into training, validation, and test sets provided with each dataset and applied the same evaluation protocol.

4.1. Gender Classification

In this section, we report the results for gender classification on the KHATT, QUWI, and HHD datasets. Table 1 lists the total amount of images in each of them. We added to Table 1 a ‘majority probability’ column because the datasets are imbalanced. The majority probability indicates the random guess accuracy.

Table 1. The total amount of training, validation, and testing images in each dataset. The ‘majority probability’ column describes the imbalance in the dataset (percentage of the samples from larger class).

Dataset	Images	Male	Female	Majority Probability (%)	Train	Val	Test
KHATT	2000	1354	666	67.7	1400	300	300
QUWI	1900	884	1016	53.5	1128	-	772
HHD	702	351	351	50	560	72	70

4.1.1. The Results on the KHATT Dataset

The KHATT dataset is provided with its official split: training (70%), validation (15%), and testing (15%) sets. For the experiment, we utilized the full dataset—2000 paragraph images written by 1000 writers. The results are presented in Table 2. We found only one additional study [33] that utilized the *full* KHATT dataset—1000 writers and 2000 paragraph images. Additional works [9,15–17,34,49] utilized only small subsets of KHATT with 75, 100, or 165 writes, and used segmented text line images, and their results are not directly comparable. We can see that B-ResNet achieves 76.17% classification accuracy—an increase by 2% in comparison to the results of [33] and much above the majority probability (majority probability indicates the results of the random guess).

Table 2. Gender prediction accuracy on the KHATT dataset.

Method	Accuracy (%)	Study
B-ResNet	76.17	proposed
ATP-DenseNet	74.1	Xue et al. [33]

4.1.2. The Results on the ICDAR 2013 Dataset

Table 3 compares the results obtained for gender classification with the results reported in the literature. There are three classification configurations: training and testing on documents written in English, training and testing on documents written in Arabic, and training and testing on documents in both, English and Arabic.

We can see that the results of the B-ResNet are among the top-two results in all scenarios, being the winner in the two mono-script schemes and second place in the multi-script task. English and Arabic handwriting have very different patterns, and it can be expected the multi-script results will be lower than the mono-script. In all the cases, the achieved results are much above the random guess.

Table 3. Comparison of classification accuracy (%) for gender classification on the ICDAR 2013 dataset.

Method	En/En	Ar/Ar	En+Ar/En+Ar	Study
B-ResNet	88.33	85.23	78.04	proposed
SVM	-	-	63.6	Tan et al. [23]
SVM	75.5	77.7	77.8	Akbari et al. [12]
ANN	69	71.9	79.3	Akbari et al. [12]
SVM	77.98	76.17	75	Gattal et al. [18]
SVM	-	-	66.3	Bi et al. [14]
SVM, LR, kNN (ensemble)	-	-	65.71	Maken and Gupta [20]
ATP-DenseNet	-	-	71.8	Xue et al. [33]
CNN	75	74	71	Rabaev et al. [8]
<i>Top ICDAR results</i> (different systems)	79	74	76	retrieved from [12]

4.1.3. The Results on the HHD Dataset

The HHD dataset consists of 702 document samples: 351 males and 351 females. We used the official split from [8]. Table 4 presents the results of the B-ResNet and compares them to the top results reported in [8]. We can see that the results of B-ResNet are on par with the results of [8], which all are far above the random guess. The size of the HHD is much smaller than that of the KHATT and QUWI, and the B-ResNet model has a much higher number of parameters (around 21M) in comparison to other models. Such a large number of parameters require a large dataset to train the model. We believe that training on a larger dataset will improve the B-ResNet results further.

Table 4. Gender prediction accuracy on the HHD dataset.

Method	Accuracy (%)	Study
B-ResNet	84	proposed
Xception	85	Rabaev et al. [8]
EfficientNet	84	Rabaev et al. [8]
NashNet	84	Rabaev et al. [8]

4.2. Age Classification

In this part, we report the results of age classification on the KHATT dataset. To the best of our knowledge, KHATT is the only available public offline dataset that contains age labels. Similar to the gender classification, we used the official split provided with the dataset. Age labels are provided in the form of four classes: (1) “>15”, (2) “16–25”, (3) “26–50”, (4) “>50”. The major class is the age group between 15 and 25, which constitutes 64.4% of the entire dataset. Table 5 presents the KHATT split.

Comparison of results is difficult due to the fact that previous studies utilized different settings—a much smaller number of writers utilized in the experiments, text line segmented images as opposed to paragraph-level images, and a different number of age classes.

The experimental settings in the previous studies were as follows. In the studies of Basavaraja et al. [39] and Bouadjeneq et al. [16], only two age ranges were utilized: “16–25 years” and “26–50 years”, 135 samples for each age range with the split 2/3 : 1/3 for training and testing sets. Bouadjeneq et al. [15] collected 135 samples for each of the three classes: “under 15 years”, “16–25 years”, and “26–50 years”. For each class, 2/3 of the samples were used for the training step, while the remaining 1/3 were used for testing the system. All of the aforementioned studies employed segmented text line images.

Our experiments were performed on the *full* KHATT dataset with 1000 writers and 2000 *paragraph* images, classified into four ages group, as provided by the official split. We did not find any previous study that utilized the full KHATT dataset for age classification. The average accuracy on the KHATT dataset is 66.65%. The results for age classification are lower than the results for gender classification. Age classification seems to be a harder task as compared to gender classification, where only two classes are given in later. While inspecting the confusion matrix presented in Table 6, we can notice that the lowest results are for the age groups “<15” and “>50”. In the age group “<15”, only one image out of 42 was classified correctly, and in the group “>50” no image was assigned to the correct

group. We attribute such low results to the fact that these two age groups are under-represented, especially the “>50” group—only 34 out of 2000 paragraphs are written by people above 50 years old.

Table 7 compares the results of different studies. Recall that previous studies used very different experimental settings—a much smaller number of writers, segmented line images as opposed to full paragraph images, and only two or three age classes—which does not make it possible to perform a fair quantitative comparison with the results of the proposed method. To compare the B-ResNet performance with the systems presented in [15,16,39], we calculated the accuracy for only two and three age classes. The B-ResNet archives the second-best result for two age classes scenario and the best result for three classes scenario, outperforming the result of [15] by almost 12%. It should be noted that in all the cases, our model was trained to classify four classes. We believe that training our model on two or three groups (instead of all four as we did) would improve its performance for those groups.

Table 5. The KHATT split (number of images) for age classification problem. The major class is the age group between 15 and 25, which constitutes 64.4% of the entire dataset.

Age	Train	Test	Val
<15	172	42	38
16–25	894	196	198
26–50	310	56	60
>50	24	6	4

Table 6. The confusion matrix for age classification on the KHATT dataset.

	<15	16–25	26–50	>50
<15	1	37	4	0
16–25	0	168	26	0
26–50	0	29	29	0
>50	0	3	3	0

Table 7. Comparison of accuracy scores for age classification accuracy on the KHATT dataset for different numbers of age classes.

Method	Acc (%)	Study	Settings
Two age classes			
Hu moments and k-means clustering	64.44	[39]	line images, 270 writers in total
SVM & Fuzzy MIN-MAX rules combination	81.11	[16]	line images, 270 writers in total
B-ResNet	78.17	proposed	paragraph images, the full dataset
Three age classes			
SVM	55.55	[15]	line images, 405 writers in total
B-ResNet	67.30	proposed	paragraph images, the full dataset
Four age classes			
B-ResNet	66.65	proposed	paragraph images, the full dataset

5. Conclusions and Future Work

This work focuses on automatic gender and age classification from handwriting. To the best of our knowledge, our work is the first attempt to apply B-ResNet—the B-CNN architecture, where ResNet is instantiated into both parallel CNN blocks—to a handwriting analysis with the purpose of gender and age classification. Moreover, this is the first work on age classification using a deep learning model. We demonstrate that, in most cases, the performance of the B-ResNet performance is superior to that of other models. In the future, we plan to incorporate the attention mechanism into the B-ResNet model. Experiments in [50] show that the attention mechanism significantly reduces the classification error on fine-grained datasets of general images. Additionally, we would like to explore the correlation between gender and age classification problems, e.g., whether age can be more accurately classified inside a certain gender group and vice versa.

Author Contributions: Conceptualization, I.R. and M.L.; methodology, I.R. and M.L.; software, I.A. and O.W.; validation, I.A. and O.W.; formal analysis, I.R. and M.L.; investigation, I.R. and M.L.; resources, I.A. and O.W.; writing—original draft preparation, I.R. and M.L.; writing—review and editing, I.R. and M.L.; supervision, I.R. and M.L.; project administration, I.A. and O.W.; All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

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

References

1. Huber, R.A.; Headrick, A.M. *Handwriting Identification: Facts and Fundamentals*; CRC Press: Boca Raton, FL, USA, 1999. [\[CrossRef\]](#)
2. Hamid, S.; Loewenthal, K.M. Inferring gender from handwriting in Urdu and English. *J. Soc. Psychol.* **1996**, *136*, 778–782. [\[CrossRef\]](#)
3. Upadhyay, S.; Singh, J.; Shukla, S. Determination of Sex Through Handwriting Characteristics. *Int. J. Curr. Res. Rev.* **2017**, *9*, 11. [\[CrossRef\]](#)
4. Hardyck, C.; Goldman, R.; Petrinovich, L. Handedness and sex, race, and age. *Hum. Biol.* **1975**, *47*, 369–375.
5. Vlachos, F.; Bonoti, F. Explaining age and sex differences in children’s handwriting: A neurobiological approach. *Eur. J. Dev. Psychol.* **2006**, *3*, 113–123. [\[CrossRef\]](#)
6. Marzinotto, G.; Rosales, J.C.; El-Yacoubi, M.A.; Garcia-Salicetti, S.; Kahindo, C.; Kerhervé, H.; Cristancho-Lacroix, V.; Rigaud, A.S. Age-related evolution patterns in online handwriting. *Comput. Math. Methods Med.* **2016**, *2016*, 1–15. [\[CrossRef\]](#)
7. Lin, T.Y.; RoyChowdhury, A.; Maji, S. Bilinear convolutional neural networks for fine-grained visual recognition. *IEEE Trans. Pattern Anal. Mach. Intell.* **2017**, *40*, 1309–1322. [\[CrossRef\]](#)
8. Rabaev, I.; Litvak, M.; Asulin, S.; Tabibi, O.H. Automatic Gender Classification from Handwritten Images: A Case Study. In Proceedings of the International Conference on Computer Analysis of Images and Patterns, Online, 28–29 September 2021; pp. 329–339. [\[CrossRef\]](#)
9. Navya, B.; Shivakumara, P.; Shwetha, G.; Roy, S.; Guru, D.; Pal, U.; Lu, T. Adaptive multi-gradient kernels for handwriting based gender identification. In Proceedings of the International Conference on Frontiers in Handwriting Recognition, ICFHR, Niagara Falls, NY, USA, 5–8 August 2018; pp. 392–397. [\[CrossRef\]](#)
10. Navya, B.; Swetha, G.C.; Shivakumara, P.; Roy, S.; Guru, D.S.; Pal, U.; Lu, T. Multi-Gradient Directional Features for Gender Identification. In Proceedings of the 2018 24th International Conference on Pattern Recognition (ICPR), Beijing, China, 20–24 August 2018; pp. 3657–3662. [\[CrossRef\]](#)
11. Ahmed, M.; Rasool, A.G.; Afzal, H.; Siddiqi, I. Improving handwriting based gender classification using ensemble classifiers. *Expert Syst. Appl.* **2017**, *85*, 158–168. [\[CrossRef\]](#)
12. Akbari, Y.; Nouri, K.; Sadri, J.; Djeddi, C.; Siddiqi, I. Wavelet-based gender detection on off-line handwritten documents using probabilistic finite state automata. *Image Vis. Comput.* **2017**, *59*, 17–30. [\[CrossRef\]](#)
13. Alaei, F.; Alaei, A. Gender Detection Based on Spatial Pyramid Matching. In Proceedings of the International Conference on Document Analysis and Recognition, Lausanne, Switzerland, 5–10 September 2021; Volume 12824 LNCS; pp. 305–317. [\[CrossRef\]](#)
14. Bi, N.; Suen, C.Y.; Nobile, N.; Tan, J. A multi-feature selection approach for gender identification of handwriting based on kernel mutual information. *Pattern Recognit. Lett.* **2019**, *121*, 123–132. [\[CrossRef\]](#)
15. Bouadjeneq, N.; Nemmour, H.; Chibani, Y. Age, gender and handedness prediction from handwriting using gradient features. In Proceedings of the 2015 13th International Conference on Document Analysis and Recognition (ICDAR), Nancy, France, 23–26 August 2015; pp. 1116–1120. [\[CrossRef\]](#)
16. Bouadjeneq, N.; Nemmour, H.; Chibani, Y. Robust soft-biometrics prediction from off-line handwriting analysis. *Appl. Soft Comput.* **2016**, *46*, 980–990. [\[CrossRef\]](#)
17. Bouadjeneq, N.; Nemmour, H.; Chibani, Y. Fuzzy integrals for combining multiple SVM and histogram features for writer’s gender prediction. *IET Biom.* **2017**, *6*, 429–437. [\[CrossRef\]](#)
18. Gattal, A.; Djeddi, C.; Siddiqi, I.; Chibani, Y. Gender classification from offline multi-script handwriting images using oriented Basic Image Features (oBIFs). *Expert Syst. Appl.* **2018**, *99*, 155–167. [\[CrossRef\]](#)
19. Likforman-Sulem, L.; Cordasco, G.; Esposito, A. Is On-Line Handwriting Gender-Sensitive? What Tells us a Combination of Statistical and Machine Learning Approaches. In Proceedings of the International Conference on Pattern Recognition and Artificial Intelligence, Paris, France, 1–3 June 2022; pp. 287–298. [\[CrossRef\]](#)
20. Maken, P.; Gupta, A. A method for automatic classification of gender based on text-independent handwriting. *Multimed. Tools Appl.* **2021**, *80*, 4573–24602. [\[CrossRef\]](#)

21. Moetesum, M.; Siddiqi, I.; Djeddi, C.; Hannad, Y.; Al-Maadeed, S. Data driven feature extraction for gender classification using multi-script handwritten texts. In Proceedings of the International Conference on Frontiers in Handwriting Recognition, ICFHR, Niagara Falls, NY, USA, 5–8 August 2018; pp. 564–569. [[CrossRef](#)]
22. Siddiqi, I.; Djeddi, C.; Raza, A.; Souici-meslati, L. Automatic analysis of handwriting for gender classification. *Pattern Anal. Appl.* **2015**, *18*, 887–899. [[CrossRef](#)]
23. Tan, J.; Bi, N.; Suen, C.Y.; Nobile, N. Multi-feature Selection of Handwriting for Gender Identification Using Mutual Information. In Proceedings of the 2016 15th International Conference on Frontiers in Handwriting Recognition (ICFHR), Shenzhen, China, 23–26 October 2016; pp. 578–583. [[CrossRef](#)]
24. Topaloglu, M.; Ekmekci, S. Gender detection and identifying one’s handwriting with handwriting analysis. *Expert Syst. Appl.* **2017**, *79*, 236–243. [[CrossRef](#)]
25. Mirza, A.; Moetesum, M.; Siddiqi, I.; Djeddi, C. Gender classification from offline handwriting images using textural features. In Proceedings of the 2016 15th International Conference on Frontiers in Handwriting Recognition (ICFHR), Shenzhen, China, 23–26 October 2016; pp. 395–398. [[CrossRef](#)]
26. Altman, N.S. An introduction to kernel and nearest-neighbor nonparametric regression. *Am. Stat.* **1992**, *46*, 175–185. [[CrossRef](#)]
27. Hassaine, A.; Al Maadeed, S.; Aljaam, J.; Jaoua, A. ICDAR 2013 competition on gender prediction from handwriting. In Proceedings of the International Conference on Document Analysis and Recognition, Washington, DC, USA, 25–28 August 2013; pp. 1417–1421. [[CrossRef](#)]
28. Djeddi, C.; Al-Maadeed, S.; Gattal, A.; Siddiqi, I.; Souici-Meslati, L.; Abed, H.E. ICDAR2015 competition on Multi-script Writer Identification and Gender Classification using ‘QUWI’ Database. In Proceedings of the International Conference on Document Analysis and Recognition, ICDAR, Nancy, France, 23–26 August 2015; pp. 1191–1195. [[CrossRef](#)]
29. Djeddi, C.; Al-Maadeed, S.; Gattal, A.; Siddiqi, I.; Ennaji, A.; El Abed, H. ICFHR2016 Competition on multi-script writer demographics classification using “QUWI” database. In Proceedings of the 2016 15th International Conference on Frontiers in Handwriting Recognition (ICFHR), Shenzhen, China, 23–26 October 2016; pp. 602–606. [[CrossRef](#)]
30. Gornale, S.S.; Kumar, S.; Patil, A.; Hiremath, P.S. Behavioral biometric data analysis for gender classification using feature fusion and machine learning. *Front. Robot. AI* **2021**, *8*, 685966. [[CrossRef](#)]
31. Dargan, S.; Kumar, M.; Tuteja, S. PCA-based gender classification system using hybridization of features and classification techniques. *Soft Comput.* **2021**, *25*, 15281–15295. [[CrossRef](#)]
32. Illouz, E.; David, E.O.; Netanyahu, N.S. Handwriting-based gender classification using end-to-end deep neural networks. In Proceedings of the International Conference on Artificial Neural Networks, Rhodes, Greece, 4–7 October 2018; pp. 613–621. [[CrossRef](#)]
33. Xue, G.; Liu, S.; Gong, D.; Ma, Y. ATP-DenseNet: A hybrid deep learning-based gender identification of handwriting. *Neural Comput. Appl.* **2021**, *33*, 4611–4622. [[CrossRef](#)]
34. Rahmanian, M.; Shayegan, M.A. Handwriting-based gender and handedness classification using convolutional neural networks. *Multimed. Tools Appl.* **2021**, *80*, 35341–35364. [[CrossRef](#)]
35. Bouadjenek, N.; Nemmour, H.; Chibani, Y. Fuzzy integral for combining SVM-based handwritten soft-biometrics prediction. In Proceedings of the 2016 12th IAPR Workshop on Document Analysis Systems (DAS), Santorini, Greece, 11–14 April 2016; pp. 311–316.
36. Al Emran, M.; Naief, S.; Hossain, M. Handwritten Character Recognition and Prediction of Age, Gender and Handedness Using Machine Learning. Ph.D. Thesis, BRAC University, Dhaka, Bangladesh, 2018.
37. Upadhyay, S.; Singh, J. Estimation of age through handwriting characteristics in female writers. *Indian J. Sci. Technol.* **2017**, *10*, 1–8. [[CrossRef](#)]
38. Zouaoui, F.; Bouadjenek, N.; Nemmour, H.; Chibani, Y. Co-training approach for improving age range prediction from handwritten text. In Proceedings of the 2017 5th International Conference on Electrical Engineering-Boumerdes (ICEE-B), Boumerdes, Algeria, 29–31 October 2017; pp. 1–5.
39. Basavaraja, V.; Shivakumara, P.; Guru, D.S.; Pal, U.; Lu, T.; Blumenstein, M. Age Estimation using Disconnectedness Features in Handwriting. In Proceedings of the 2019 International Conference on Document Analysis and Recognition (ICDAR), Sydney, Australia, 20–25 September 2019; pp. 1131–1136. [[CrossRef](#)]
40. Al-Showarah, S.A. Dynamic recognition for user age-group classification using hand-writing based finger on smartphones. In Proceedings of the 2019 10th International Conference on Information and Communication Systems (ICICS), Irbid, Jordan, 11–13 June 2019; pp. 140–146.
41. Shin, J.; Maniruzzaman, M.; Uchida, Y.; Hasan, M.; Mehedi, A.; Megumi, A.; Suzuki, A.; Yasumura, A. Important Features Selection and Classification of Adult and Child from Handwriting Using Machine Learning Methods. *Appl. Sci.* **2022**, *12*, 5256. [[CrossRef](#)]
42. Lin, T.Y.; RoyChowdhury, A.; Maji, S. Bilinear CNN models for fine-grained visual recognition. In Proceedings of the IEEE International Conference on Computer Vision, Santiago, Chile, 7–13 December 2015; pp. 1449–1457.
43. Simonyan, K.; Zisserman, A. Very deep convolutional networks for large-scale image recognition. *arXiv* **2014**, arXiv:1409.1556.
44. He, K.; Zhang, X.; Ren, S.; Sun, J. Deep residual learning for image recognition. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Las Vegas, NV, USA, 27–30 June 2016; pp. 770–778.

45. Mahmoud, S.A.; Ahmad, I.; Alshayeb, M.; Al-Khatib, W.G.; Parvez, M.T.; Fink, G.A.; Märgner, V.; El Abed, H. KHATT: Arabic offline handwritten text database. In Proceedings of the 2012 International Conference on Frontiers in Handwriting Recognition, Bari, Italy, 18–20 September 2012; pp. 449–454. [[CrossRef](#)]
46. Mahmoud, S.A.; Ahmad, I.; Al-Khatib, W.G.; Alshayeb, M.; Parvez, M.T.; Märgner, V.; Fink, G.A. KHATT: An open Arabic offline handwritten text database. *Pattern Recognit.* **2014**, *47*, 1096–1112. [[CrossRef](#)]
47. Al Maadeed, S.; Ayouby, W.; Hassaine, A.; Aljaam, J.M. QUWI: An Arabic and English handwriting dataset for offline writer identification. In Proceedings of the 2012 International Conference on Frontiers in Handwriting Recognition, Bari, Italy, 18–20 September 2012; pp. 746–751. [[CrossRef](#)]
48. Rabaev, I.; Kurar Barakat, B.; Churkin, A.; El-Sana, J. The HHD Dataset. In Proceedings of the International Conference on Frontiers in Handwriting Recognition, Dortmund, Germany, 8–10 September 2020; pp. 228–233.
49. Morera, Á.; Sánchez, Á.; Vélez, J.F.; Moreno, A.B. Gender and Handedness Prediction from Offline Handwriting Using Convolutional Neural Networks. *Complexity* **2018**, *2018*, 1–15. [[CrossRef](#)]
50. Eshratifar, A.E.; Eigen, D.; Gormish, M.; Pedram, M. Coarse2Fine: A two-stage training method for fine-grained visual classification. *Mach. Vis. Appl.* **2021**, *32*, 1–9. [[CrossRef](#)]