Electronic Preprint for Journal of Information Processing Vol.30

Regular Paper

Construction of Japanese Imperial Diet Database Using Deep Neural Network

Naoki Nonaka,[a)] Naoto Nonaka[1,b)]

Received: September 5, 2021, Accepted: December 26, 2021

Abstract: In this work we construct a database of the Japanese Imperial Diet. The Imperial Diet was established in 1890 after the promulgation of the Constitution of the Empire of Japan and its historical analysis is crucial to understand the actual functioning of the Japanese Diet. Since the minutes of the Imperial Diet were publicly available only in image format, textization is an imperative process for further analysis. Following the recent advancement of the deep neural networks (DNNs), especially in the character recognition, we apply DNNs to construct a text database of the Imperial Diet. In the course of textization, we trained DNNs using multiple datasets while introducing a novel approach of applying separate batch normalization to datasets. The results of the tentative analysis show a significant potential to deepen our understanding of the development of parliamentary democracy in Japan.

Keywords: OCR, parliament minutes, deep neural network

1. Introduction

The Imperial Diet started in 1890 just after the promulgation of the Constitution of the Empire of Japan in 1889, which brought a European-style constitutional monarchy to Japan. Its establishment marked a critical turning point for Japan's modernization effort. Under the imperial constitution, a bicameral system constituted of the "Shuugi-in (literally the House of the People)" and the "Kizoku-in (House of Peers)". Although both Houses had roughly equal authorities, "Shuugi-in" possessed an crucial predominance in deciding the matters on budget. In fact, as demonstrated in the emergence of the party governments in the 1920's, this power on the budget made the "Shuugi-in" the main pillar of the government at that time. We can regard this development as the precursor of the democratic development in the years after the war. In addition, as the recent research development has shown clearly, post-war Japan's Diet inherited largely the main mechanism, rules and even basic principles of pre-war Imperial Diet. Therefore, if we should understand the actual irregularities of Japan's Diet now, understanding the Imperial Diet, especially the "Shuugi-in", its Lower House, is essential.

In recent years, as the technology of natural language processing has become more widespread and improved, there have been reports of studies that have analyzed the text of parliamentary proceedings in the field of political science. Proksh and Slapin [20] analyze the speech in the European Parliament, and Quin [21] applies a topic model to analyze the speech activities in the Senate of the US Congress. As will be discussed later, there are several more recent studies in this field. In addition to the researches on the analysis of parliamentary speech texts, efforts are also being made to develop a database of parliamentary proceedings data. By combining the recently developed text analysis methods with conventional statistical ones, we can now expect enormous possibilities of new research which would enable advanced quantitative evaluation of the activities in parliaments as well as gaining new knowledge in political science in general.

However, in order to perform text analysis, the proceedings must be in a machine-readable format. The proceedings of parliaments in European countries or post-war Japanese Diet are already available in text format. Thus, collection and analysis of these parliamentary proceedings are relatively easy. On the other hand, the proceedings of the pre-war Imperial Diet are available only as image data. Therefore, to analyze the pre-war Imperial Diet, we first need to extract text information from the image data.

Recently, deep neural networks (DNNs) have shown significant performance in various fields, such as image classification, natural language processing and speech recognition [4], [6], [10], [11], [27]. DNNs have also been applied to detect characters in natural images [24], [28]. The success of DNNs in these applications shows the powerful potential of DNNs in text extraction from image data. Thus, in this work which aims to construct the speech text database, we apply DNNs to image data of the Imperial Diet to extract its text information.

The paper is organized as follows. First, we describe related works on parliamentary proceedings data and the recognition of text from images using DNNs. Next, we describe the data used to convert the proceedings into text and the dataset used to train the model. Subsequently, we explain the preprocessing of the minutes data and the training conditions for the DNN model in Section 4. Following Section 4, we show the results of experiments on the training conditions of the DNN model and a brief analysis based on the constructed database. Finally, we discuss limitations and possible future works in Section 7.

[1] Gakushuuin University, Toshima, Tokyo 171–8588, Japan
[a)] naokinonaka1989@gmail.com
[b)] 19950068@gakushuin.ac.jp

2. Related Work

2.1 Text Analysis of Parliamentary Speech

In the field of political science, parliamentary research with abundant data sources has entered a new stage in recent years. At first, analytical techniques such as Wordscore and Wordfish (which later evolved into Wordshaol) were introduced to quantitative analysis on parliamentary minutes. For example, Ref. [20] use Wordfish to analyze the speeches in the European Parliament focusing on the roles of political parties. As an analysis of debate contents, a topic model was applied to parliamentary minutes in the US Senate [21]. Around this time, various research developments have been seen [7], [12], [19], [23], [26]. These recent works are the demonstrations of not only technical innovations but also significant advancements in the fields of comparative parliamentary studies. An overall review of recent works on parliamentary texts is provided by Ref. [1].

In addition to the advancement of quantitative analysis methods, the structurization of parliamentary minute data is also underway. In Europe, for example, a multilateral cooperation framework has been organized and large-scale efforts are underway for years. One is CLARIN (https://www.clarin.eu/), which is promoting the digitization of linguistic information in a form that covers the humanities and social sciences. The other is PolMine (https://polmine.github.io/), which is building a database of deliberation records in parliament in each country. Both are based on the EU framework, with many countries participating in it. In particular, the latter is specialized in the parliamentary texts, and the research utilizing it is being actively promoted. In addition, ParlaCRARIN workshops are held regularly, and the results of those workshops are being accumulated (Refs. [8], [9]).

As for the data on Japanese Diet, basic information is provided by the National Diet Library via its internet service. It also provides the API for searching speeches and speakers within the post-war Diet activities. Based on this, himawari corpus and KokkaiGiinHakusho (https://kokkai.sugawarataku.net/) are known for data structurization. The himawari corpus provides XML formatted parliamentary minutes of the plenary session and budget committee of post-war Shuugi-in until 2012. KokkaiGiinHakusho provides web service to search for various activities of parliamentarians including their speech. However, all of these deal with data from the Japanese Diet since 1947. Therefore, one of the most salient limitations of these works is that they do not offer the basis for the studies with longer time perspectives. Even the service provided by the National Diet Library is fundamentally limited as its data format is not digitised for the period before the war. Given these situations, our efforts to construct an enlarged and digitised database on the Imperial Diet would have an important significance.

2.2 Deep Neural Networks for Character Recognition

In recent years, DNN has shown significant performance in a wide range of fields such as image classification, natural language processing and speech recognition [4], [6], [10], [11], [27]. Above all, it shows high performance in the task of image classification including character recognition. In addition to simple character classification, a DNNs has been applied in character detection and subsequent recognition in a natural image. For example, Ref. [24] applied Convolutional Recurrent Neural Network on scene text recognition tasks and Ref. [28] used attention based models to extract information from street view images. Also, DNNs have been applied to detect non alphabetical characters [22]. These results show the promising results of DNNs in optical character recognition (OCR).

In the context of OCR applied to historical Japanese text, there are two lines of research. One is an application to Kuzushiji, which was commonly used before the Meiji era. As an attempt to recognize Kuzushiji, Ref. [5] applied DNNs to recognize Kuzushiji. Another line of research is recognition of printed text after the Meiji era. Kindai-OCR [15] is a general-purpose model to apply OCR on modern Japanese text. To handle various formats, they take an approach of character recognition after detecting the position of a character based on the CRAFT approach [3]. In addition to Kindai-OCR, Ref. [13] applies a CRAFT approach to develop a character segmentation model for printed text, using the proceedings of the Imperial Diet. While the CRAFT approach for detecting characters is one of the promising methods, in this study we focused on improving the accuracy of a DNN model for character classification. As a result of improvement in classification accuracy, we have constructed a database that can be used for substantive political science analysis.

3. Data

In this section, we describe the data collection of the minutes and the data set used to train the DNN model.

3.1 Minute Data of the Imperial Diet

The image data of the minutes was obtained from the Imperial Diet minutes search system operated by the National Diet Library (https://teikokugikai-i.ndl.go.jp/#/). The search system covers the entire period under the Japanese Imperial Constitution, from the first session that began in 1890 to the end of the 22nd session in 1947. The proceedings of both the "Kizoku-in (House of Peers)" and the "Shuugi-in (literally the House of the People)" are open to the public, not only for plenary sessions but also for each committee. In addition, the information of speakers recorded in the minutes is digitized and linked in the search system, and it is possible to search for a speaker or to check the list of speakers in the minutes. However, except for the minutes of the 22nd legislative term held after World War II, these minutes are published as image data, and therefore, it is not possible to retrieve the contents of the speeches as text data.

In this research, we extract text information from minutes of the Imperial Diet published as image data. The extraction of text was conducted for a total of 18 years, from the beginning of the 15th legislative term in 1924 to the end of the 20th legislative term in 1942. The 15th legislative term corresponds to the age of the "Party Governments", which is important for understanding the development of democracy in Japan. The 20th legislative term includes the period up to the beginning of the Pacific War, which is important for understanding the process of Japan's entry into the war. We targeted the data of the Shuugi-in, which is equiva-

Table 1 Summary of three datasets used to train character classification model. We selected characters which appeared in the transcribed minute of the 22nd legislative term's minute from each data source.

	Number of characters	Total number of samples
Minute	2,308	249,416
ETL-9	2,443	488,600
Font	3,387	361,644

lent to the Lower Houses in other countries, which have the direct confidence relation with the government, thus the decisive political power. As the plenary sessions are critically important for the final decisions of the Diet, we target in this paper the proceedings of the plenary.

3.2 Dataset Used to Train DNN Model

In order to train a DNN model to transcribe the minutes of the Imperial Parliament, we used three datasets. The first is a dataset of manually annotated minutes of the Imperial Parliament. The second dataset is the ETL-9 dataset, which consists of handwritten JIS I level kanji and hiragana characters. The third dataset is a dataset obtained from font data. These datasets were jointly used to train the DNN model. The details of each dataset are shown in **Table 1** and processing procedures are described below.

3.2.1 Minute Dataset

First, as a dataset for training DNN models to recognize text in the minutes, we manually annotated a part of the minutes. We selected 40 pages from 15,122 pages of the proceedings of the Shuugi-in from the 15th to the 20th legislative terms and manually annotated them. The annotated data was made into a square by cutting out the area of each character and padding it with white. Then, all the characters were resized to a size of 40×40 pixels. Finally, the image was binarized by Otsu's method [18]. The resulting dataset (henceforth referred to as Minute dataset) contained 2,308 different characters, for a total of 249,416 samples.

The Minute dataset is suitable for training DNN models, as it is derived from real minutes, however it has some limitations for training DNN models. Annotation of minutes data is costly and is expensive to obtain large scale data. Thus, it may be insufficient for training DNNs that require large amounts of data. In addition, the distribution of characters contained in the obtained data set is non-uniform because it follows the frequency of appearance in the minutes. Therefore, in this study, two additional datasets were used to compensate for the lack of data and the non-uniformity of appearance of each character.

3.2.2 ETL-9 Dataset

In order to compensate for the imbalance of characters and insufficient amount of data in the Minute dataset, we used the ETL-9 dataset (http://etlcdb.db.aist.go.jp/specification-of-etl-9). The ETL-9 dataset is a handwritten character dataset consisting of 3,036 characters, including 2,965 JIS I Kanji characters and Hiragana characters. The dataset consists of 607,200 samples created by 4,000 authors. In this study, we trained the model using the ETL-9B dataset, which is pre-binarized by Otsu method [18]. Each image was resized to a size of 40×40 pixels as in the Minute dataset. The characters used for training DNN were 2,443 characters from the Minute dataset and the textualized minutes of the 90th to 92nd Imperial Parliament, out of the 3,036 characters in the ETL-9 dataset.

3.2.3 Font Dataset

For training the DNN model, we used a dataset with font data in addition to the ETL-9 dataset. Since the kanji in ETL-9 are new style characters, the old style characters used in the Imperial Diet minutes are not included in the dataset. In order to compensate for the lack of the data of the old style characters, we constructed a dataset by obtaining the image data of the old characters from the font data. The font data used in our work is a collection of publicly available font data, including styles such as Mincho and Gothic. To construct dataset for training character recognition model, we first obtained character images from the collected font data. Next, the obtained character images were padded to make them square and resized to 40×40 pixels as in the Minute dataset. Finally, the dataset was binarized using the Otsu method [18]. The resulting dataset (hereafter referred to as the Font dataset) contained 3,387 different characters, for a total of 3,616,644 samples.

4. Method

In this section we explain preprocessing of data and training procedure of DNN model.

4.1 Preprocessing

In this section, we describe the preprocessing we applied on the collected images of the minutes. The minutes of Imperial Parliament are published as a government gazette and have a certain format. Therefore, we used a rule-based method to divide them into lines based on the structure of the format.

The outline of the preprocessing procedure is shown in **Fig. 1**. First, the image was processed to remove the surrounding white space. After applying Gaussian blurring to the image, the pixel values were added horizontally to calculate a histogram. Based on the calculated histogram, the upper and lower white areas were removed. Subsequently, the same process was applied in the vertical direction, and the white areas on the left and right edges were removed.

After removing surrounding white space, the paragraphs in each page were cropped. Similar to the first step procedure, we first applied Gaussian blurring to the images. After blurring, the histograms were obtained by adding up the pixel values in the horizontal direction. Since there are four paragraphs in each page of the target document, we divided the document into four paragraphs based on the histogram.

Following the above procedure, we extracted a line from each paragraph. First, Gaussian blurring was applied to the image of each paragraph, and the pixel values were added together in the vertical direction to obtain a histogram. The rows were cropped based on the histogram values.

Finally, each cut line was divided into characters. Based on the histogram obtained by adding the values of the horizontal pixels, the image was divided into areas corresponding to one character. White pixels were added to the top, bottom, left and right of each character to make the image square. The resulting image was resized to 40×40 pixels, the same size as the dataset used

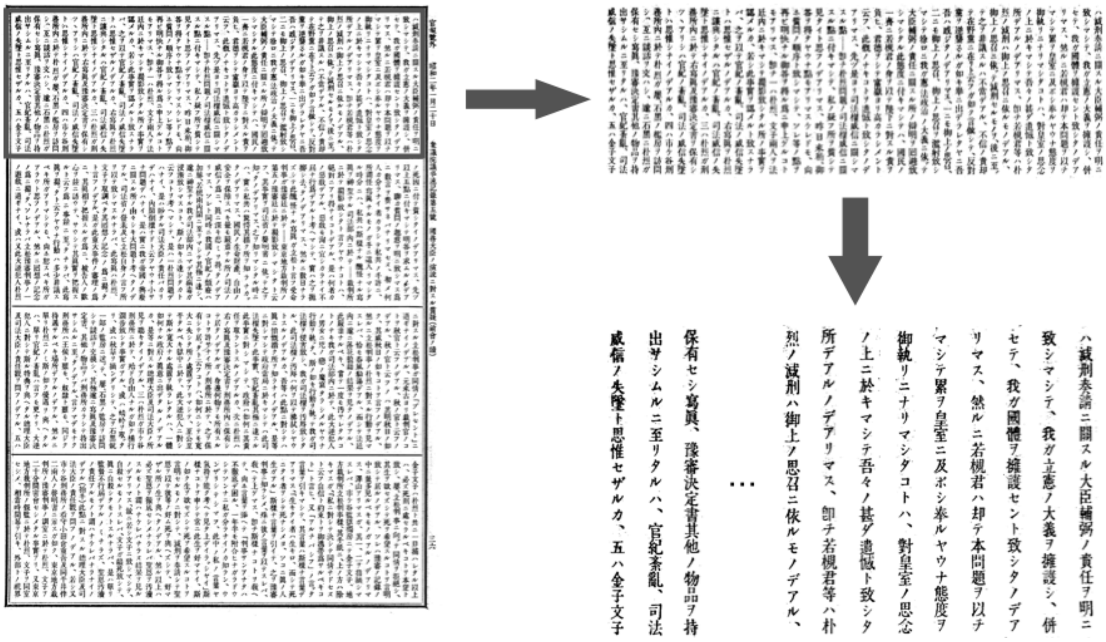
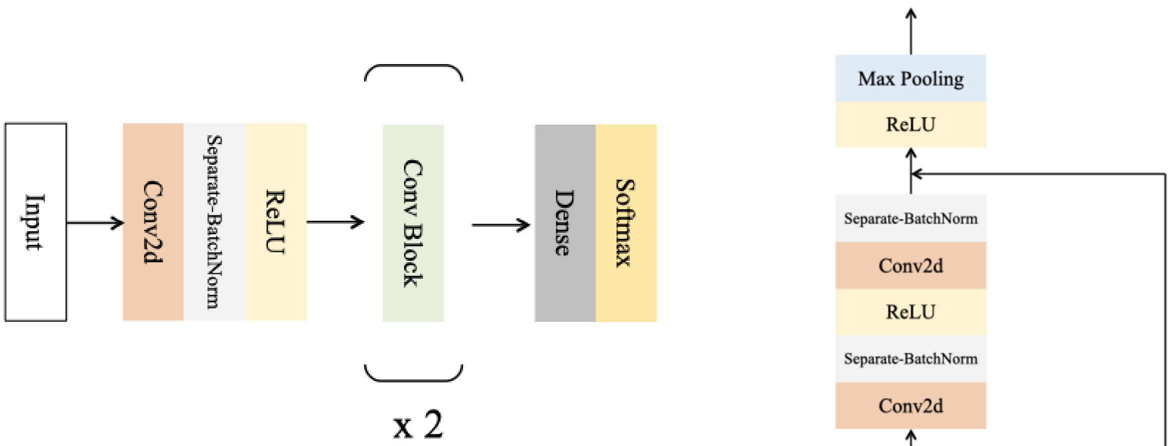


Fig. 1 Overview of preprocessing of collected minute image data. First we cut out columns using a histogram of horizontally accumulated pixel values. Subsequently, for each column, we cut out rows using a histogram of vertically accumulated pixel values. Finally, after splitting into rows, we split the row image into individual characters by using a histogram of horizontally accumulated pixel values.



(a) Overall architecture

(b) ConvBlock

Fig. 2 Model architecture of character classification model. (a) shows overall architecture. Convblock structure shown in (b) is stacked 2 times in the model. We use ReLU as an activation function and use separate batch normalization which applies separate batch normalization based on source of training data.

for training, and given to the DNN model.

4.2 Classification Model

As a DNN model to classify characters, we adopted a convolutional neural network. The network takes a 40×40 pixel, two dimensional image data as an input, and outputs a character prediction for each input image. The overall structure of the network is shown in **Fig. 2** (a). Our model consists of two ConvBlocks, shown as in Fig. 2 (b), and followed by a fully connected layer and a softmax activation. Each ConvBlock has two convolutional layers with filter length of 4, and each convolution layer has 32 filters. Max pooling layer in ConvBlock is applied in every block, beginning from the first ConvBlock.

The three datasets described in Section 3 were split into train, validation and test set respectively. The data for each character in each dataset was divided into train, validation and test sets in the ratio of 8 : 1 : 1. We trained the DNN model using a train set, selected optimal parameters using a validation set and obtained the final result using a test set.

We applied a separate batch normalization approach described below, and inserted a rectified linear unit, ReLU after every batch normalization layer. The final fully connected layer and softmax activation outputs a vector of length 3,391, which corresponds to the number of characters.

4.2.1 Separate Batch Normalization

In our character classification model, three datasets are trained simultaneously. These datasets are from different domains such as letterpress, handwriting, and font, and the ratio of characters in each dataset is also different. Batch normalization [14] is a method that speeds up learning by calculating the mean and variance of each mini-batch, and also has the effect of regularization. However, applying the same batch normalization to data from different domains may have a negative effect. In the context of training image classification models with adversarial samples, Ref. [29] have shown by applying auxiliary batch normalization to adversarial samples improves classification accuracy. Therefore, in this study, we use separate batch normalization, a method that applies different batch normalization to each dataset.

4.2.2 Hyper Parameter Setting

The details of the training setting is as follows. We used AdamW optimizer [16] to train the models with a learning rate of 0.00001. The batch size was set to 1,024. We trained models for 250 epochs and monitored validation loss every 5 epochs. After training we selected the best model based on validation set loss and report test set loss and accuracy. We trained each model five times and reported average test set loss and accuracy as a final result. To mitigate the class imbalance problem, the ratio of samples belonging to each class to the total number of data was calculated, and the reciprocal of the ratio was used as the weight for each class.

4.3 Post-processing

In this subsection, we explain the post-processing for the result of transcription by DNN model. The result of transcription includes misrecognized characters due to the error of character segmentation in preprocessing, misclassification by DNN model, and misdetection of non-character regions. Therefore, to improve the quality of transcribed results by DNN model, we apply following procedure.

First, we used Google Cloud Vision (https://cloud.google.com/vision) to correct the recognition results. Google Cloud Vision is a publicly available service that provides image recognition and character reading functions. In this study, we applied Google Cloud Vision to the images of each paragraph obtained from the preprocessing, and read the characters. Subsequently, for each paragraph, we aligned the reading results of Google Cloud Vision with the results of our DNN model. After alignment, if the two recognition results did not match, the recognition result with higher probability of occurrence was selected based on the bigram. The bigram was created based on the co-occurrence frequency of characters in the text of 90-92 Imperial Diet proceedings. In case of successive discrepancies in recognition results, the results were corrected using bigram in order from the front.

Following the above procedure, we applied a rule-based correction of the frequent misrecognition results. There are three main patterns of correction. The first one is the misrecognition of characters that contain spaces or part of the line dividing paragraphs at the beginning and end of the line, and the DNN model is applied to them. The second is a case where a single character is split into multiple characters as a result of excessive splitting.

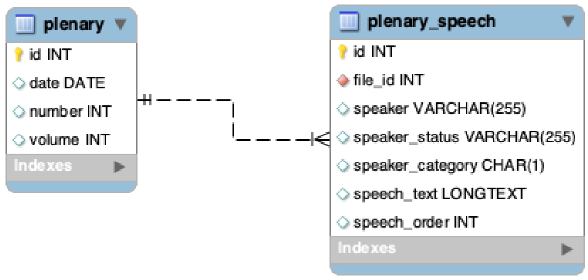


Fig. 3 ER diagram of created database schema. Each legislative term has a single schema. Each schema consists of two tables, “plenary” and “plenary_speech”.

For example, the letter “吾” is split into “五” and “口”, and the letter “シ” is split into “、” and “ン”. The third correction was the integration of symbols. For example, “(”, “[”, and “c” were unified into “(”.

Finally, the names and status of the speakers have been corrected. The names and status of speakers in a parliamentary session are crucial for political science analysis. In the minutes, the beginning of each speaker’s statement is indicated by a “○” sign, followed by the speaker’s name and title. Therefore, we checked the lines with “○” at the beginning of the line and manually corrected the names and titles of the speakers after “○”. After applying all post-processing procedures, we stored the resulting text into the database.

4.4 Structure of Database

After post-processing, the obtained text was stored into a database. The database was created for each legislative term. Based on the information of the speaker identified in the text, a series of statements by one speaker was stored in the database as one record. The details of the database structure are described below.

As shown in the ER diagram of the created database in **Fig. 3**, there are two tables for each database. The “plenary” table holds the information of each volume of the minutes. In addition to the unique id, this table contains the date and time of each plenary session, the order of the session, and the volume for each session held during the relevant legislative term. On the other hand, the “plenary_speech” table records the speeches in each volume. The “file_id” corresponds to the “id” in the “plenary” table, so that it is possible to retrieve the speeches by specifying a specific volume. Each record in the “plenary_speech” table contains the name of the speaker, the title of the speaker, the category of the title, the text of the speech, and the order of the speakers. The created database can reconstruct the information of the minutes in the official gazetteer by using both plenary and “plenary_speech” tables.

5. Result of Transcription Process

In this section we discuss the result of the DNN model.

5.1 Effect of Separate Batch Normalization

First, we examined the effect of separate batch normalization used in our DNN model.

Batch normalization is a method that calculates the mean and

Table 2 Comparison of separate batch normalization and regular batch normalization. We report the result on a test set of manually annotated imperial minute dataset. As a result of five independent experiments, we observed improvement both in average loss and average accuracy.

	Average loss ↓	Average accuracy ↑
Batch Normalization	0.194	0.943
Separate Batch Normalization	0.142	0.986

Table 3 Evaluation of jointly using multiple datasets. We report the result on a test set of manually annotated imperial minute dataset. As a result of five independent experiments, we observed best test set loss when jointly using all three datasets.

Minute	ETL-9	Font	Average loss ↓	Average accuracy ↑
✓	-	-	0.264	0.988
✓	✓	-	0.213	0.985
✓	-	✓	0.148	0.987
✓	✓	✓	0.142	0.986

variance of each mini-batch during training, and normalizes the samples using the moving average during testing. In general, batch normalization has the advantage of obtaining regularization and stabilizing the training process. However, in this study, we used three datasets with different sources at the same time, and applying the same batch normalization to all the data may have a negative effect on training. Therefore, we applied separate batch normalization, which calculates the mean and variance for each of the three datasets used and applies normalization to them.

The result of comparing standard batch normalization and separate batch normalization is shown in **Table 2**. In each condition, five independent training sessions were conducted and the test set loss was recorded. We compared the average loss and accuracy of the five training runs, and found that the results of separate batch normalization outperformed those of standard batch normalization. This result suggests the effectiveness of applying separate batch normalization when jointly using multiple datasets to train DNN models.

In the following experiments, we used the model with separate batch normalization.

5.2 Effect of Jointly Using Multiple Dataset for Training

Subsequently, we examined the impact of using multiple datasets. In this study, we used three datasets from different domains. The Minute dataset has a high coverage rate of characters used in the minute, including old style kanji, but it has a high degree of imbalance in the included samples. The ETL-9 dataset contains a large number of samples for a wide range of kanji characters, but the coverage of characters is low since it does not include old style kanji. Therefore, we examined the classification performance of the DNN model when the Minute dataset is combined with the ETL-9 and Font datasets.

Five independent experiments were conducted in each setting, and the average loss and accuracy of the test set on the Minute dataset are shown in **Table 3**. Lowest average loss was achieved when all three datasets, Minute, ETL-9, and Font, were jointly used during training. When only one of the datasets, ETL-9 or Font, was used in addition to the Minute dataset, the degree of decrease in the average loss was greater when the Font dataset was added. These results indicate that the use of the two datasets, ETL-9 and Font, especially the Font dataset, in addition to the Minute dataset improves the performance of the model.

Table 4 Accuracy of character recognition. Accuracy was evaluated using manually annotated imperial minute dataset. For each setting, we calculated alignment score between ground truth and predicted characters and divided by lengths of ground truth text. By combining the results of the DNN model and Google Cloud Vision, we observed the best score.

	Average score ↑
DNN model	0.925
Google Cloud Vision	0.774
Post-processed	0.929

5.3 Accuracy of Character Recognition

Next, we evaluated the results of the DNN model trained on the three datasets simultaneously, as well as the results of the post-processing. For the evaluation of the results, we used 40 pages of manually annotated minutes of the Imperial Diet. We calculated the score for each prediction result using the following procedure. The alignment scores between the annotated results and the predicted results were calculated using dynamic programming. The obtained scores were divided by the total number of characters on each page to calculate the score for each page, and the average of the scores for each page was used as the final score. We evaluated the results in three cases: using the prediction results of the DNN model, using the results of Google Cloud Vision, and results after post-processing.

The results are shown in **Table 4**. The results of transcription by our DNN model showed relatively high accuracy, although there were cases where the system recognized numbers and some non-character regions at the beginning and end of lines as characters. On the other hand, Google Cloud Vision, as a general-purpose character recognition system, recognized some characters as alphabets or symbols, which are not included in the minutes of the Imperial Diet, resulting in lower accuracy. Since the tendency of misclassification was different between our DNN model and the Google Cloud vision results, we combined these results through post-processing. Further details of character recognition results are shown in Appendix A.1. As a result of post-processing, we observe slight improvement over the result of our DNN model. Thus, we stored post-processed results into our database.

6. Analysis of Database

In this section we show basic analysis of the obtained database to demonstrate its usefulness. We use the legislative term as the unit of analysis for the period segmentation. One legislative term is the period between a dissolution and the next dissolution [*1].

6.1 Frequently Appearing Words

First, we calculated the number of characters in the database we created. **Table 5** shows the number of characters per legisla-

[*1] The choice of the unit of time period naturally depends on the target of analysis and the method used. The legislative term is fairly broadly used. See for example, Refs. [2], [3], or [30]. When organizing basic data on parliaments or conducting comparative parliamentary analysis, the period from one dissolution to the next is often treated as a unit of analysis.

Table 5 Total number of characters in each legislative term (LEG stands for legislative term).

	LEG15	LEG16	LEG17	LEG18	LEG19	LEG20	Total
Number of characters	8,854,273	3,110,088	2,889,709	7,070,029	3,471,277	9,349,955	34,745,331
Term	1924 – 1928	1928 – 1930	1930 – 1932	1932 – 1936	1936 – 1937	1937 – 1942	1924 – 1942
Length of term	3 year 7 month	1 year 9 month	1 year 8 month	3 year 10 month	10 month	4 year 8 month	16 year+
Avg. number of characters	205,913	148,099	144,485	153,696	347,128	166,963	-

Fig. 4 Word cloud based on created database. We used all speech texts from the 15th to 20th legislative term. Prominent words such as "帝国議会" (Imperial Diet), "開院"(opening of the House) and "奉答" (reply to the throne) are specific to the pre-war Imperial Diet.

tive term and the length of each term. The total number of characters for each term divided by the length of the term (number of months) was similar except for the 19th legislative term. These figures were roughly identical to the average number of characters in the Shuugi-in in the early postwar period.

Next, we analyzed the words in the database. Neologd was used as a dictionary and MeCab was used for word segmentation. Since the morphological analyzer is designed for modern Japanese, the accuracy of the segmentation is low when it is applied directly to the proceedings of the Imperial Diet, which uses old characters and katakana. Therefore, we converted old-style kanji into the corresponding new-style, and converted all katakana into hiragana before conducting morphological analysis. Words that were classified as nouns or adjectives by the morphological analysis were included in the analysis. However, words classified as personal names, words formed with a single letter, and words considered to be misclassified by the transcription model were excluded. The analysis also excluded technical words, which are used in the same way both in pre-war and post-war periods in the course of proceedings. They include, for example,"議長" (Speaker of the House), "委員長" (Chairperson of committees), "法案" (bill), "採決" (vote).

Figure 4 shows a word cloud created using the frequently occurring words. The results show that many words such as "帝国議会" (Imperial Diet), "奉答" (reply to the throne), "開院"(opening of the House), "陛下" (His Majesty), and "摂政" (regent) appear in the word cloud. These words are considered to reflect the basic nature of the Imperial Diet as the supporting organ to the Emperor, and are not found in the postwar Diet proceedings.

We extracted words characteristic of each legislative term using the term frequency-inverse document frequency (TF-IDF) method. As in the analysis of the number of words, words classified as personal names, words formed by a single letter, and words considered to be misclassified by the transcription model were excluded from the analysis. Since the purpose of the TF-IDF method is to extract words that are characteristic of each particular legislative term, words that appear in more than five legislative terms out of 6 legislative terms (from 15th to 20th) were excluded from the analysis. The top 20 words extracted from each legislative term by the TF-IDF method are shown in **Table 6**.

Table 6 shows the words that are characteristic of each legislative term, suggesting the issues that were mainly discussed during each particular term. In the 15th and 16th legislative terms, which correspond roughly to the period of the party governments, many words related to the electoral system, such as "小選挙区" (Small-sized electoral district), "中選挙区" (Medium-sized electoral district) and "一人区" (single - seat constituency) are included. It also includes words related to labor-related laws and criminal law, which were deeply associated with the introduction of the male universal franchise. It can be said that, during the 15th

Table 6 Top 20 words of TF-IDF value in each legislative term. In the 15th and 16th legislative terms, words related to the electoral system such as “小選挙区” (Small-sized electoral district) and “中選挙区” (Medium-sized electoral district) are included. On the other hand, words related to war and military affairs, such as “支那事変” (Sino-Japanese War), “戦時体制” (wartime order) are included in the 19th and 20th legislative terms.

LEG	Words
15	高等師範学校，政友本党，擧区，凡平，妥协，減刑，富力，編輯，両て，復興局，農務省， 正誤，松島遊廓，革新倶楽部，査問，労働争議調停法，清浦，地租条例，加藤内閣，恩赦
16	擧区，両税委譲，小選挙区，田中首相，両て，小選挙区制，一人区，済南事件，妥協， 山東出兵，信愛，中選挙区，入党，沖島，三土，中選挙区制，妥当，誘揚，地租条例，再保険
17	抵当証券，産業合理化，自動車道，租年，刑事補償法，安保，愛用，軍令部長，内閣総理大臣臨時代理， 霧社事件，通用門，補助艦，抵当証券法，告発，国際決済銀行，保税，計量士，船員保険，電気事業者，労働争議調停法
18	満洲国，赤字，国民同盟，いんふれ，臨時利得税，助産師，傷痍軍人，自力更生，速記士， 教護，公傷，為替手形，長野県諏訪，会商，冷害，統制経済，北見国，振出，補整，日本製鉄株式会社
19	恩給金庫，防空，結城，満洲国，国民健康保険，揮発油，赤字，税制改革，米穀自治管理法，保健所，いんふれ， 台湾拓殖株式会社，帝国燃料興業株式会社，医療費，国民健康保険組合，自動車製造，商工組合中央金庫，輸入組合，新潟県北，国際情勢
20	支那事変，恩給金庫，厚生省，臨時利得税，厚生大臣，大東亜，銃後，日本発送電株式会社， 青年学校，満洲国，将兵，防空，いんふれ，戦時体制，営団，摩擦，法人税，物品税，完遂，税制改革

and 16th legislative terms, with the introduction of the single constituency electoral system and then the men’s universal suffrage, the importance of the Imperial Diet, especially the House of Representatives, gradually increased. Frequent appearance of these words related to the electoral system clearly reflected this political trend. Obviously these views are not novel. However, the important thing is that we were able to quantitatively confirm what has been discussed qualitatively through traditional historical research methods *2. In the 17th and 18th Diets, we can find words related to the economy, such as “抵当証券” (mortgage-backed securities), “産業合理化” (industrial rationalization), “国際決済銀行” (Bank for International Settlements), “赤字” (deficit), and “いんふれ” (inflation). Probably these words reflect the especially difficult economic situation in the age of the Great Depression. In the 19th Diet and later, words related to diplomacy and military affairs such as “満州国” (Manchuria) and “防空” (air defense) appear, and words related to energy policy such as “帝国燃料興業株式会社” (Imperial Fuel Industries) and “揮発油” (gasoline) are also included. In the 20th Diet, words that directly reflect the wartime situation, such as “支那事変” (Sino-Japanese War), “戦時体制” (wartime order), “完遂” (completion), “将兵” (generals and soldiers), and “大東亜” (Greater East Asia) ranked high.

6.2 Speech Volume Per Speaker

Following the analysis of words contained in the constructed database, we conducted an analysis of the speakers. First, we calculated the total number of times people spoke in each legislative term and the number of people who spoke. The top three speakers in each legislative term were also calculated and displayed. The results are shown in **Table 7**. In each legislative term, the Speaker of the House has the highest number of statements, followed by the Deputy Speaker. The Speakers and the Deputy Speakers have the highest number of speeches because they make very frequent interventions regarding the procedural management. The third most frequent speaker is the particular member who is charged with making procedural motions for swift advancement of deliberations.

*2 See for example Refs. [17], and [25] as the historical studies on the Imperial Diet of this period.

Table 7 Total number of speeches recorded, total number of speakers and top three speakers in terms of speech count of each legislative term. For all terms, the top two speakers are the House Speaker and the Deputy Speaker.

LEG	Total number of speech	Number of speakers	Top 3 speakers
15	13,617	392	粕谷義三 小泉又次郎 作間耕逸
16	6,165	247	元田肇 清瀬一郎 川原茂輔
17	4,511	215	藤澤幾之輔 小山松壽 作田高太郎
18	9,502	361	秋田清 濱田國松 青木雷三郎
19	3,971	287	富田幸次郎 岡田忠彦 松永東
20	11,484	507	小山松壽 服部崎市 金光庸夫

Table 8 Total speech length per person for each legislative term. For each legislative term the top speaker was the House Speaker. Prime Minister (Osachi Hamaguchi) or minister of state (such as Korekiyo Takahashi) is also included as the second or the third.

LEG	Top1	Top2	Top 3
15	粕谷義三 1,386,354	小泉又次郎 632,612	濱口雄幸 235,820
16	元田肇 298,476	川原茂輔 258,069	清瀬一郎 188,225
17	藤澤幾之輔 516,427	小山松壽 89,943	高橋熊次郎 76,849
18	秋田清 1,602,488	濱田國松 469,833	高橋是清 189,232
19	富田幸次郎 687,432	岡田忠彦 266,107	河原田稼吉 58,499
20	小山松壽 1,449,865	富田幸次郎 532,922	田子一民 315,617

Next, we conducted an analysis of the amount of statements made by each person. The total number of characters in each person’s speech text was calculated for each legislative term. **Table 8**

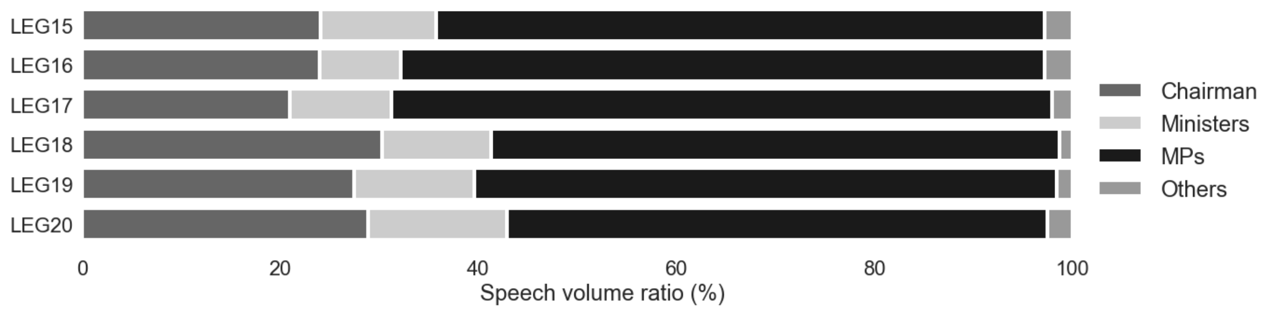


Fig. 5 Speech volume ratios of speaker categories of each legislative term. We observe a decrease of MPs' speech ratio after the 18th legislative term.

shows the top three speakers and the total number of characters in each legislative term. The Speaker of the House has the most number of characters in each parliamentary term. However, there is a possibility that the Speaker's statement includes the list of law names, the list of voting results, and the contents read by the clerk, which are not actually spoken by the Speaker. This may result in an apparent increase in the number of words. In the second and third positions of the total number of words, the Prime Minister and Ministers of State such as Osachi Hamaguchi and Korekiyo Takahashi were included. As for these important government people, although their number of speeches are rather limited, the total numbers of characters are large, reflecting their relative weights in the debates.

Subsequently, we calculated the amount of statements per legislative term, divided by speaker categories. The categories of speakers and their speaking percentages are important factors in analyzing the debate structure of parliaments. The categories are Speaker and Deputy Speaker, Ministers of State including the Prime Minister, Members of Parliament, and Others. The results of calculating the percentage of each category are shown in **Fig. 5**. The percentage of statements made by the Speaker and Vice Speaker was 21.0% in the 17th legislative term, the lowest, and 30.3% in the 18th, the highest. The percentage of statements by the Minister of State was 8.2% in the 16th legislative term, the lowest, and 14.0% in the 20th, the highest. The percentage of statements by parliamentarians was the lowest in the 20th legislative term at 54.6%, and the highest in the 16th at 65.0%. The percentage of statements made by parliamentarians was above 60% in the 15th, 16th, and 17th legislative terms, while it was below 60% in the 18th, 19th, and 20th. This trend may suggest that as time neared the war period, activity level of parliamentarians fell, reflecting the general tendency of the loss of speech freedom in Japanese society as a whole.

7. Discussion and Conclusion

In this study, we have constructed a database of the proceedings of the Imperial Diet, which was effectively one of the most important organs of the pre-war constitutional monarchy of Japan. The constructed database covers proceedings of the plenary sessions of the Shuugi-in, the Lower House, from 15th to 20th legislative term. The first two legislative terms, 15th and 16th, correspond roughly to the period of "Party Government", a crucial period in light of the development of Japanese democracy. The 20th legislative term, on the other hand, corresponds to the period during which Japan began the Pacific War. The analysis of parliamentary texts from these periods is, therefore, an important element in understanding the history of the parliamentary democracy in Japan.

Since the proceedings of the pre-war Imperial Diet are publicly available only in image format, we have developed a DNN model to extract text information from image data. To train the DNN model for text extraction, we devised two points. First, we trained the model by jointly using three datasets with different characteristics. Second, we developed a separate batch normalization method, which applies batch normalization separately to three datasets used during training. The result of experiments showed that simultaneous use of three datasets combined with the use of separate batch normalization improved the classification performance of our DNN model. We should note that choice of font styles used during the training may affect the accuracy of the DNN model, leaving room for further improvements by suitable selection of font styles.

To further improve the transcription accuracy, we post-processed the results obtained from our DNN model. We applied Google Cloud Vision, a publicly available text recognition system, to the proceedings and then aligned the results of it with those of our DNN model. The experimental results showed, combining results of our DNN model and Google Cloud Vision, together with a rule-based post-processing, improves the overall accuracy.

Although our approach enabled the text transcription of the proceedings, the constructed database has several limitations. The major limitations fall into four points. First, there are some failures of histogram-based paragraph segmentation. This was mainly caused by distortions and dirt of the scanned data. Second, there are several misclassifications of character. This was caused by both errors of character segmentation from row and misclassification by our DNN model. Third, resulting arrangements of sentences may differ in some cases, such as a table being posted across multiple paragraphs. Lastly, speech and non-speech, such as voting result and explanatory texts of the bill themselves, are not separated properly. Since voting results and summaries of bills are added later mostly in the position after chairman's speech, the volume of chairman's speech may be over-calculated. The first, second and third limitations may be mitigated by applying a more advanced approach on segmentation and localization.

In addition to the technical limitations mentioned above, we should remember that the coverage itself of the proceedings is limited. Beyond the data collected in this study, there are minutes of committees other than plenary sessions, those of the Kizoku-in (House of Peers), and those Shuugi-in proceedings of the periods before the 15th legislative term. For the future, we plan to convert these minutes into text so that everyone can analyze the activities of the Japanese Diet since the Meiji period.

Although a tentative one, an analysis of the proceedings in the database showed quite clearly the changes in the activities in the Diet from the period of "Party Government" to that of the Pacific War. In particular, we observed a shift in the frequency of words used in plenary sessions, reflecting the changing focus in the debates in the Diet. In the Imperial Diet, parliamentarians spoke, debated on the themes of the day. Deeply changed themes and wordings from democracy related terms like elections to the ones directly related to the war were depicted vividly in this way. By using other analytical methods such as topic models and combining them with qualitative studies of the Imperial Diet, we believe that we can contribute in a novel way to the analysis of the development of parliamentary politics in Japan.

Acknowledgments This work was supported by JSPS KAKENHI Grant Number JP20K20509.

References

[1] Abercrombie, G. and Batista-Navarro, R.: Sentiment and position-taking analysis of parliamentary debates: A systematic literature review, *Journal of Computational Social Science*, pp.1–26 (2020).

[2] Arter, D.: *Comparing and classifying legislatures*, Routledge (2013).

[3] Baek, Y., Lee, B., Han, D., Yun, S. and Lee, H.: Character Region Awareness for Text Detection, *Proc. IEEE Conference on Computer Vision and Pattern Recognition*, pp.9365–9374 (2019).

[4] Bahdanau, D., Cho, K. and Bengio, Y.: Neural machine translation by jointly learning to align and translate, arXiv preprint arXiv:1409.0473 (2014).

[5] Clanuwat, T., Lamb, A. and Kitamoto, A.: Kuronet: Pre-modern Japanese kuzushiji character recognition with deep learning, *2019 International Conference on Document Analysis and Recognition* (*ICDAR*), pp.607–614, IEEE (2019).

[6] Devlin, J., Chang, M.-W., Lee, K. and Toutanova, K.: Bert: Pre-training of deep bidirectional transformers for language understanding, arXiv preprint arXiv:1810.04805 (2018).

[7] Fernandes, J.M., Goplerud, M. and Won, M.: Legislative Bellwethers: The Role of Committee Membership in Parliamentary Debate, *Legislative Studies Quarterly*, Vol.44, No.2, pp.307–343 (2019).

[8] Fišer, D., Eskevich, M. and de Jong, F. (Eds.): *Proceedings of the Second ParlaCLARIN Workshop*, European Language Resources Association (2020).

[9] Fišer, D., Eskevich, M. and de Jong, F. (Eds.): *Proc. LREC2018 Workshop ParlaCLARIN: Creating and Using Parliamentary Corpora*, European Language Resources Association (ELRA) (2018).

[10] Graves, A., Mohamed, A.-R. and Hinton, G.: Speech recognition with deep recurrent neural networks, *2013 IEEE International Conference on Acoustics, Speech and Signal Processing*, pp.6645–6649, IEEE (2013).

[11] He, K., Zhang, X., Ren, S. and Sun, J.: Deep residual learning for image recognition, *Proc. IEEE Conference on Computer Vision and Pattern Recognition*, pp.770–778 (2016).

[12] Høyland, B. and Søyland, M.G.: Electoral reform and parliamentary debates, *Legislative Studies Quarterly*, Vol.44, No.4, pp.593–615 (2019).

[13] Iida, S., Takemoto, Y., Ishikawa, Y., Takata, M. and Joe, K.: A Study of Text Detection Methods for Early-Modern Japanese Books, *Mathmatical Modeling and Problem Solving* (*MPS*), Vol.2021, No.4, pp.1–6 (2021).

[14] Ioffe, S. and Szegedy, C.: Batch normalization: Accelerating deep network training by reducing internal covariate shift, *International Conference on Machine Learning*, pp.448–456, PMLR (2015).

[15] Le, A.D., Mochihashi, D., Masuda, K., Mima, H. and Ly, N.T.: Recognition of Japanese historical text lines by an attention-based encoder-decoder and text line generation, *Proc. 5th International Workshop on Historical Document Imaging and Processing*, pp.37–41 (2019).

[16] Loshchilov, I. and Hutter, F.: Decoupled weight decay regularization, arXiv preprint arXiv:1711.05101 (2017).

[17] Murai, R.: 政党内閣制の成立 *1918〜27* 年 (*Establishment of the Party Cabinet System*), Yuuhikaku (2005).

[18] Otsu, N.: A threshold selection method from gray-level histograms, *IEEE Trans. Systems, Man, and Cybernetics*, Vol.9, No.1, pp.62–66 (1979).

[19] Proksch, S.-O., Lowe, W., Wäckerle, J. and Soroka, S.: Multilingual sentiment analysis: A new approach to measuring conflict in legislative speeches, *Legislative Studies Quarterly*, Vol.44, No.1, pp.97–131 (2019).

[20] Proksch, S.-O. and Slapin, J.B.: Position taking in European Parliament speeches, *British Journal of Political Science*, Vol.40, No.3, pp.587–611 (2010).

[21] Quinn, K.M., Monroe, B.L., Colaresi, M., Crespin, M.H. and Radev, D.R.: How to analyze political attention with minimal assumptions and costs, *American Journal of Political Science*, Vol.54, No.1, pp.209–228 (2010).

[22] Ren, X., Zhou, Y., He, J., Chen, K., Yang, X. and Sun, J.: A convolutional neural network-based chinese text detection algorithm via text structure modeling, *IEEE Trans. Multimedia*, Vol.19, No.3, pp.506–518 (2016).

[23] Rheault, L. and Cochrane, C.: Word embeddings for the analysis of ideological placement in parliamentary corpora, *Political Analysis*, Vol.28, No.1, pp.112–133 (2020).

[24] Shi, B., Bai, X. and Yao, C.: An end-to-end trainable neural network for image-based sequence recognition and its application to scene text recognition, *IEEE Trans. Pattern Analysis and Machine Intelligence*, Vol.39, No.11, pp.2298–2304 (2016).

[25] Shuuggin and Sangiin: 議会制度百年史：帝国議会史下巻 (*Hundred Years of the National Diet: Imperial Diet ii*), Vol.4, Ministry of Finance (1990).

[26] Slapin, J.B. and Kirkland, J.H.: The sound of rebellion: Voting dissent and legislative speech in the uk house of commons, *Legislative Studies Quarterly*, Vol.45, No.2, pp.153–176 (2020).

[27] Szegedy, C., Liu, W., Jia, Y., Sermanet, P., Reed, S., Anguelov, D., Erhan, D., Vanhoucke, V. and Rabinovich, A.: Going deeper with convolutions, *Proc. IEEE Conference on Computer Vision and Pattern Recognition*, pp.1–9 (2015).

[28] Wojna, Z., Gorban, A.N., Lee, D.-S., Murphy, K., Yu, Q., Li, Y. and Ibarz, J.: Attention-based extraction of structured information from street view imagery, *2017 14th IAPR International Conference on Document Analysis and Recognition* (*ICDAR*), Vol.1, pp.844–850, IEEE (2017).

[29] Xie, C., Tan, M., Gong, B., Wang, J., Yuille, A.L. and Le, Q.V.: Adversarial examples improve image recognition, *Proc. IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp.819–828 (2020).

[30] Zubek, R.: Negative agenda control and executive–legislative relations in east central Europe, 1997–2008, *The Journal of Legislative Studies*, Vol.17, No.2, pp.172–192 (2011).

Appendix

A.1 Example of Text Recognition

From the results of the recognition process, we selected the parts that differed from the data used for training, compared them with the transcriptions from actual minutes, and calculated the recognition accuracy of the characters. Three speeches by the prime minister were used for the calculation of recognition accuracy: a speech by Takaaki Kato on July 1, 1924, a speech by Reijiro Wakatsuki on January 18, 1927, and a speech by Yoshikazu Tanaka on January 28, 1928. Speech texts were obtained from https://worldjpn.grips.ac.jp/. Numerical evaluations showed that the character recognition accuracies of the speeches by Kato, Wakatsuki and Tanaka were 0.913, 0.946 and 0.925, respectively. This result is almost consistent with the accuracy of 0.928 after Post-processing shown in Table 4. As an example of the actual recognition results, **Fig. A·1** shows the comparison of the character recognition results of the beginning of Kato's speech with the image of the minutes.

○國務大臣(子爵加藤高明君) 諸君、私ハ
今回不敏ヲ以チマシテ敢テ自ラ揣ラズ、內
閣組織ノ大命ヲ拜シマシテ、過ル六月十一
日御親任ヲ辱フ致シ、茲ニ諸君ト相見ユル
ニ至リマシタノハ、私ノ洵ニ光榮トスル所
デアリマスル、內閣成立後日尚ホ淺クゴザ
イマシテ、諸般施設ノ準備未ダ整頓ニ至リ
マセヌノハ誠ニ已ムヲ得ナイ所デアリマシ
テ、此點ハ豫メ諸君ノ御諒察ヲ請ハナケレ
バナラヌ所デアリマス、刻下內外ノ情形ヲ
審案致シマスルニ、外列國ハ力ヲ內ニ蓄ヘ
マシテ、勢ヲ外ニ伸バサントスルニ汲々ト
致シ、隨ッテ內ニ顧ミマスレバ、人心ノ緊張
ヲ要スル甚ダ切ナルモノガアリマス、隨ッテ
庶政百般ニ涉リマシテ更新スベキモノガ頗
ル多々ナリト認ムルノデアリマス、今右ノ
中ニ就キマシテ緊要ナルモノ數項ヲ擧ゲ、
所信ノ一端ヲ披瀝致シマシテ、御參考ニ資
シタイト思ヒマス、諸君、締盟各國トノ交
際ハ倍々敦厚ヲ加ヘマシテ、何等渝ル所ガ
ナイノハ誠ニ慶スベキ所デアリマス、既ニ
御承知ノヤウニ過般米國ノ議會ニ於キマシ
テ新移民法ガ成立致シタノデアリマス、右
立法中、我ガ移民排斥ヲ目的ト致シマスル
條項ハ、正義公平ノ要求ニ合致致サナイモ
ノデアリマシテ、之ヲ阻止セントスル我ガ
朝野ノ熱心ナル努力ガアリマシタルニ拘ラ
ズ、斯ル立法ガ遂ニ實施トナリマシタノハ
頗ル遺憾トスル所デアリマス、政府ハ飽マ
デ合理的手段ニ依リ、本問題ノ解決ニ努メ
タイト思フノデアリマス、露國ニ對スル親
交ノ回復スルニ至ランコトハ政府ノ固ヨリ

諸君、私ハ
今回不敏ヲ以チマシテ敢テ自ラ揣ラス、内
閣組織ノ大命ヲ拝シマシテ、過ル六月十三
日御親任ヲ辱フ致シ、茲ニ諸君ト相見ユル
ニ至リマシタノハ、私洵ニ光栄ト致ス所
三テアリマスル、内閣成立後日尚ホ浅クコサ
イマシテ、諸般施設ノ準備未タ整頓ニ至リ
マセヌハ誠ニ已ムヲ得ナイ所テアリマシ三
テ、此点ハ、予メ諸君ノ御諒察ヲハナケレ
ハナラヌ所テテリマス、二刻下内外ノ情形ヲ
三案致シマスルニ、外列国ハ力ヲ内ニ三
マシテ、勢ヲ外ニ伸ハサントスルニ汲々ト
欲シテ内顧ミマスレハ、人心ノ緊張
字ヲ要スル甚タ切ナルモノカアリマス、随ツテ
等庶政百般ニ渉リマシテ更新スモノ
ルト認ムルノテアリマス、今右ノ
中ニ就キマシテ緊要ナルモノ数項ヲ挙ケ
所信ノ一端ヲ披瀝致シマシテ、御参考ニ資
シタイト思ス、諸君、締盟各国ノ交
一際ハ厚ヲ加ヘマシテ、何等ル所カ
三ナイノハ誠ニスヘキ所テアリマス、既ニ
御承知ノヤウニ過般米国議会ニ於キマシ
テア移民法カ成立致シタノテアリマス、右
立法中、我カ移民排斥ヲ目的ト致シマスル
ハ、二正義公平要求ニ合致致サナイモ
ノテアリマシテ、之ヲ阻止セントスル我カ
朝野熱心ナル努力カアリマシタルニ拘ラ
ス、斯ルコ法カ途ニ益トナリマシタノハ
頗ルトスル所テアリマス、政府ハ飽マ
一テ合理的手段依リ、本問題ノ解決ニ努メ
三タイト思フノテアリマス、露国ニ関スル親
交ノ回復スルニ至ランコトハ政府ノ固ヨリ

Fig. A·1 Result of text recognition by DNN model jointly trained using Minute, ETL-9 and Font datasets. The upper row is the image of the actual proceeding, and the lower row is the recognition result of the corresponding part by DNN model. The minutes are the first part of the speech by then Prime Minister Takaaki Kato at the 49th Imperial Diet held on July 1, 1924. Note that the speaker's information ("○國務大臣（子爵加藤高明君）") in the first line is outside the scope of the recognition model.

Naoki Nonaka Born in 1989, he obtained Ph.D. in Graduate School of Engineering at the University of Tokyo in 2018. His doctoral dissertation was an analysis of time series data obtained from social media. One of his current interests is an application of deep neural networks and other machine learning techniques.

Naoto Nonaka Born in 1958, he got his Ph.D. in 1994 in International Relations at the University of Tokyo, and since 1996 he has been professor of political science at Gakushuin University. His doctoral dissertation was an empirical comparative study of the major political parties in Japan and France. In recent years, he has been working on comparative parliamentary study with a particular focus on France, Japan and the UK. As part of this effort, he has been building a database of speech texts and conducting research by using new analysis methods developed in the field of natural language processing.

(Editor in Charge: *Sho Sato*)