

# Re-Produce Published Research Paper

*(Task for up to 3 students per team, coursework, 100%)*

We recommend reading this entire document **in detail** before starting the coursework.

## Forming your Teams:

Register your team of up to 3 people (i.e. one, two, or three students) online at: <https://forms.office.com/e/cgRpMUeB1A> (only one person per team needs to do this).

Post registration, teams can split but cannot merge, to avoid copying of code or ideas.

**Each member of the team should submit an exact copy of the final submission on Blackboard by the deadline.** The report (see below) should note the full names and usernames of all members in the team as well as a signed agreement (this can be digital) that all members contributed roughly equally. All members of the group will be given the same mark.

It is up to each team to decide their best strategy to tackle this coursework, i.e. whether to divide the tasks below, or to work together on all tasks. Contributions of team members need not be explicitly stated.

**However**, by submitting a group coursework, you are implicitly acknowledging that all members of the team contributed approximately equally. If this is not the case, you should email the unit director with any issues encountered during the coursework (also see Appendix A).

Note, that in the past we have found no benefit in working as part of a group, there is no correlation between mark and group size. Keep in mind the communication overhead of working in a group compared to working solo.

## Task Brief

This assignment gives you the opportunity to appreciate the work required in replicating published research from a publicly available dataset and manuscript. It allows you to reflect on the experience of reproducing published results and potentially outperforming on your replication.

Gathering all the knowledge you acquired from the lectures and labs, read the paper below carefully and replicate the required results (Note: you are not required to re-produce all the paper's results). Feel free to take any pieces of code from the labs as a baseline, but the rest of the code should be originally yours.

## The Paper

Dieleman, Sander, and Benjamin Schrauwen. "End-to-end learning for music audio." *2014 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. IEEE, 2014.  
<https://ieeexplore.ieee.org/abstract/document/6854950/>

(Log in via institution to access the paper or email me if there's issues)

Note that our choice of paper is based on its simplicity and similarity to your labs, rather than its superior performance or exceptional novelty.

Please read the following information carefully **before** attempting the replication:

### 1. Architecture

For this coursework, you will only be asked to implement the Convolutional architecture without feature pooling (Figure 1(b) in the paper). **Failure to include the Convolutional architecture will result in 0 marks for the coursework!**

### 2. Magnatagatune Dataset

The authors train and test on the Magnatagatune dataset please use the train/test split that we have provided for you to use (see Dataset and Useful Code).

### 3. Required Results and Evaluation Metric

In replicating results, we expect you to provide the code to your results for the model using AUC with **raw audio** marked here from Table 1 in the paper with differing lengths and strides:

length	stride	AUC (spectrograms)	AUC (raw audio)
1024	1024	0.8690	0.8366
1024	512	0.8726	0.8365
512	512	0.8793	0.8386
512	256	0.8793	0.8408
256	256	0.8815	0.8487

**You should structure your code that the length and stride can be take in as hyperparameters.**

We haven't covered the AUC metric in Applied Deep Learning – this may be something that you're familiar with from Introduction to Machine Learning – but AUC stands for Area Under Curve. You might ask which curve, and this is generally a ROC curve (or Receiving Operator Characteristic) which it is in this paper.

## 4. Other details

In replicating the method, there were some implementation details left unspecified by the authors. We give a list of these below as well as some hints:

- The dataloader we provide will return samples that have been divided into 10 sub-clips with padding already handled.
- There is a missing layer within Figure 1 of the paper, which you will need to map the 200 units to the number of classes for the output. Note the paper mentions to average pool over the sub-clips before classification (“To compute tag predictions for a clip, we averaged the predictions over consecutive windows”)
- Batches are of size: [B, 10, 1, 34950] for batch, sub-clips, channel dim and num\_samples.
- The paper (like many deep learning papers) is light on information regarding hyperparameter values. We got our replication accuracy with sensible hyperparameters (i.e. you shouldn’t need a learning rate of 0.113523).
- You can compress the batch size and clips into one dimension with `torch.flatten()` and revert back with `torch.reshape()`.
- The loss used is Binary Cross Entropy (BCELoss)

## 5. Our Replicated Results

Replicating results from papers rarely produces the exact results as advertised. We have first created a PyTorch implementation of the paper using the data files and train/test split available to you. Our results can be found below:

Length	Stride	AUC (Raw Audio
256	256	0.78 +/- 0.02

As our re-implementation performs worse on the chosen splits, **these are the values we want you to reproduce**. Achieving significantly higher results than these will score bonus marks.

**Note:** the reported results are the averages of the raw accuracy on the final epoch for 5 runs (i.e., not necessarily the best epoch). We have included the margin of error to help you gauge what your average accuracy should roughly be.

## 6. Dataset and Helpful Code

To get you started and focused on training the method, we’ve prepared resources for this project which you can find here: [https://uob-my.sharepoint.com/:f:/g/personal/jc17360\\_bristol\\_ac\\_uk/EqODeCX\\_Y0hNsfl2YOe-MJQBPeuT85B\\_Bg6rKqitMgB55A?e=roFwyl](https://uob-my.sharepoint.com/:f:/g/personal/jc17360_bristol_ac_uk/EqODeCX_Y0hNsfl2YOe-MJQBPeuT85B_Bg6rKqitMgB55A?e=roFwyl)

We recommend that you copy the dataset stored on Blue Crystal 4 using the following command, but the zip is also provided above:

```
cp -r /mnt/storage/scratch/jc17360/MagnaTagATune /mnt/storage/scratch/<user>
```

where <user> is your username.

The OneDrive directory includes a README which you can refer to understand what each file contains. The dataset contains a train split and a test split. All of your final results should be produced on the test split. **Do not train your model on data from the test split – this will penalise your mark HEAVILY!**

## 7. Extensions

This year you are able to extend the base model using any extensions within the paper or outside the paper. We are not limiting the number of extensions either due to the following reasons:

- Sometimes a small extension can provide a nice increase in performance.
- Limiting students to 1 or 2 extensions caused people to implement multiple extensions anyway and choose the best one for the final submission.
- Sometimes a single extension mentioned within the paper can be a challenge (i.e. Spectrograms)

With this in mind, you will see in the mark scheme that up to 15% of the final mark can be gained via extensions. Please think carefully over extensions and the page limit of the report:

- We expect you to discuss the extension(s) theoretically and empirically, i.e. what it is, why you chose it, and how does it effect the performance.
- More extensions give you less space to provide this discussion in sufficient detail. Planning >2 extensions is **not** recommended.
- You **can** get full marks on this section with 1 well motivated extension that is shown to be beneficial, >1 extensions **is in no way required**.
- Please talk to us in the CW support sessions, on Teams, or over email regarding extensions if you are in any way unsure.

## Final Submission

Within your final submission you should submit the following:

1. Original code written in Python and PyTorch (other software/libraries will NOT be accepted) that replicates the published paper. You can use lab code from any or all group members. **We will run your code on Blue Crystal, so ensure that it compiles and runs.**
2. A report in the IEEE format (we recommend conference format) which can be found here: <https://template-selector.ieee.org/secure/templateSelector/publicationType>. The report should be no longer than **5 pages including references**. The report should include the following sections:
  - a. **Title and Team Members:** (names and usernames) in addition to an agreement that all members gave an almost equal contribution with signatures (See Appendix A). Note that for single person groups only the name/username is required.
  - b. **Introduction:** Definition of the problem addressed by Schindler et al. (in your own words!)
  - c. **Related Work:** A summary of more recent published work (i.e., after Dieleman et al. was published in 2014) attempting to address the same problem (up to 3 works).
  - d. **Dataset:** A description of the dataset used, training/test split size, and labels.

- e. **CNN Architecture (Dieleman et al.):** Describe the architecture(s) that you have recreated and all of its details.
- f. **Implementation Details:** Provide a summary of the steps you have undertaken to replicate the results, train the data, and obtain the results. Do not include pieces of code, but you can use pseudo code if you find this helpful.
- g. **Replicating Quantitative Results:** You need to provide your version of Table 1 from the paper with the corresponding rows given above.
- h. **Training Curves:** Provide the train/test loss curves and accuracy curves and comment on any over/underfitting you find within your training. These curves should be the same that you use in **Section G above** and can be directly gathered from tensorboard.
- i. **Qualitative Results:** This section should include sample cases where your method worked well and where it struggled based on your algorithm. You can present examples of the audio names with the ground truth and predicted classes. We expect 1 good example where your prediction works and up to 2 examples where your prediction can be criticised.
- j. **Improvements (if doing extensions):** In this section you should give information about your improvement(s) that you have made to the method (see the mark scheme for details). You should not provide code when describing these improvements, but you may use pseudo code if this helps your explanation. Your choice of improvements should be justified both theoretically and experimentally.
- k. **Conclusion and Future Work:** Summarise what your report contains in terms of content and achievements. Suggest future work that might extend, generalise, or improve the results in your report.

## Marking Guideline

Note: code and report will be checked for plagiarism/academic misconduct. Proven plagiarism will result in a coursework grade of 0 **for the whole team**. You may use AI during this project **following [University](#) and [Faculty Guidelines](#)** (you must be logged in to see this – view the AI guidance under academic integrity) in which all prompts must be cited. Bear in mind that ChatGPT and other language models struggle to write academic reports to a high standard of 2.1 or above. Hallucinated text will **penalise you heavily** as it shows a lack of understanding. Note that not referencing the prompts you use represents academic misconduct and poor academic practice!

### Up to 55%

To pass this assignment, you must produce original, complete (compiles and runs on BC4 using batch mode command and PyTorch) code that replicates the results in the paper. You should produce a report with sections A-F correct and satisfactory. A partial attempt at including sections G-I, K is given. Replication results are within 5% of those given by us above.

### Up to 60%

In addition to the above, sections G-I, K are complete and reflective of your understanding of the code and implementation. All sections are completed to an acceptable standard. Replication results are within 2% of those given by us.

### Up to 75%

In addition to the above, provide at least one extension to the method and Section J includes results of the extension(s) and discussion. Your extension(s) should show at least a marginal improvement (i.e., be strictly greater than your base results without the extension). All sections of the report should be completed to a very good standard with good discussion of the results and method.

### Up to 80%

In addition to the above, the report should be submittable to a B-class peer review conference or venue, i.e., it shows excellent understanding, correct and complete showcasing of the approach. Statements are concise, and any jargon cut out of implementation details. The chosen related work focuses on current state of the art for this problem. Extensive evidence of analysis, creativity, and originality in concise content presentation should be shown. Code is commented and could be easily understood and re-used by the reader.

### Up to 100%

In addition to the above, the code and report are exemplary, and could be given as an example for an attempt to replicate this published work. Improvements to the results are beyond marginal (i.e., greater than 1% of baseline performance without improvement).

## Universal Coursework Details

### Deadline

The deadline for submission of all optional unit assignments is 12:00 on Thursday 7<sup>th</sup> of December (the University discourages Friday deadlines!). Students should submit all required materials to the “Assessment, submission and feedback” section of Blackboard - it is essential that this is done on the Blackboard page related to the “With Coursework” variant of the unit.

### Time commitment

You are expected to work on both of your optional unit courseworks in the 3-week coursework period as if it were a working week in a regular job - that is 5 days a week for no more than 8 hours a day. The effort spent on the assignment for each unit should be approximately equal, being roughly equivalent to 1.5 working weeks each. It is up to you how you distribute your time and workload between the two units within those constraints.

You are strongly advised NOT to try and work excessive hours during the coursework period: this is more likely to make your health worse than to make your marks better. If you need

further pastoral/mental health support, please talk to your personal tutor, a senior tutor, or the university wellbeing service.

### **Academic Offences**

Academic offences (including submission of work that is not your own, falsification of data/evidence or the use of materials without appropriate referencing) are all taken very seriously by the University. Suspected offences will be dealt with in accordance with the University's policies and procedures. If an academic offence is suspected in your work, you will be asked to attend an interview with senior members of the school, where you will be given the opportunity to defend your work. The plagiarism panel are able to apply a range of penalties, depending on the severity of the offence. These include: requirement to resubmit work, capping of grades and the award of no mark for an element of assessment.

Further information on the university's academic integrity policy can be found below:

<https://www.bristol.ac.uk/students/support/academic-advice/academic-integrity/>

### **Extensions**

If you are unwell, or there is another reason why you are unable to meet a due date, you can request an extension, however you should plan your work so that your submission is not delayed by short-term circumstances such as minor illness.

Further information and guidance on how to request an extension can be found on the below link:

<https://www.bristol.ac.uk/students/support/academic-advice/assessment-support/request-a-coursework-extension/>

As of the 23/24 academic year the deadline for the submission of an extension request is 48 hours before the coursework submission deadline. If the extension deadline has passed, then please refer to the guidance on exceptional circumstances.

### **Exceptional circumstances**

If the completion of your assignment has been significantly disrupted by serious health conditions, personal problems, periods of quarantine, or other similar issues, you may be able to apply for consideration of exceptional circumstances (in accordance with the normal university policy and processes).

<https://www.bristol.ac.uk/students/support/academic-advice/assessment-support/exceptional-circumstances/#:~:text=Exceptional%20circumstances%20are%20unexpected%2C%20unavoidable,academic%20performance%20in%20an%20assessment>

Students should apply for consideration of exceptional circumstances as soon as possible when the problem occurs, using the following online form:

<https://www.bristol.ac.uk/request-extenuating-circumstances-form>

If your application for extenuating circumstances is successful, it is most likely that you will be required to retake the assessment of the unit at the next available opportunity.



## Appendix A: Working in a Group

- **What if I can't find members for a group?** As it is optional to work within a group it is expected that you can complete the coursework by yourself if needs be. You can ask Michael or Tilo for help finding a group member, but we will not force groups together.
- **How will individual marks be assigned to each member in the group?** All members in the group will receive the same marks – the coursework is marked regardless of group size in the first place.
- **What if one or more of my team members doesn't engage with the coursework?** We expect each member of the group to contribute equally (though this could be in different areas – maybe one student is working on the report, another on the base model, another researching and implementing an extension). If this isn't the case, you should contact the unit director ASAP. By submitting the report, it is assumed that all group members are happy with the contribution of their team members with the signed agreement.
- **What if one or more of my team members becomes ill/has extenuating circumstances and has to resit?** In this case it is expected that the remaining members of the group will finish and complete the coursework on their own. Remember, it is optional to work as a group, so have a contingency plan to have to work with fewer members if needed.
- **What is the signed agreement?** You can include the following at the beginning of your report: "We agree that all members have contributed to this project (both code and report) in an approximately equal manner" with signatures of all group members below it.
- **I have a Study Support Plan (SSP)** Talk to Michael ASAP and we can discuss how to we can tailor the coursework to your SSP.