Dublin City University
School of Computing

APPLICATION FOR APPROVAL OF AN UNDERGRADUATE OR TAUGHT MASTERS PROJECT INVOLVING **HUMAN PARTICIPANTS**

PROJECT TITLE	Explainable AI in Pathology - Concept Based Explainability for Mitotic Figure Detection in Whole Slide Images
PRINCIPAL INVESTIGATOR(S) ==The Principal Investigator is the project supervisor and s/he has primary responsibility for the project.==	Alessandra Mileo
START AND END DATE	15th Sept 2022 – 9th May 2023
STUDENT NAME(S), COURSE AND YEAR (E.G. EC4)	Adam Tegart, DS4
LEVEL OF RISK ==Please confirm that this project requires notification only==	Notification only: YES

1.   ADMINISTRATIVE DETAILS

1.1   WILL THE PROJECT BE UNDERTAKEN ON-SITE AT DUBLIN CITY UNIVERSITY?

YES or NO
YES

If NO, state details of the off-campus location – provide details of the approval to gain access to that location in section 2.7.

N/A

DECLARATION BY PRINCIPAL INVESTIGATOR / SUPERVISOR

The information contained herein is, to the best of my knowledge and belief, accurate. I have read the University's current research ethics guidelines, and accept responsibility for the conduct of the procedures set out in the attached application in accordance with the form guidelines, the REC guidelines (https://www.dcu.ie/researchsupport/researchethics.shtml), the University's policy on Conflict of Interest, Code of Good Research Practice and any other condition laid down by the Dublin City University Research Ethics Committee. I have attempted to identify all risks related to the project that may arise in conducting this project and acknowledge my obligations and the rights of the participants.

If there exists any affiliation or financial interest for researcher(s) in this project or its outcomes or any other circumstances which might represent a perceived, potential or actual conflict of interest this should be declared in accordance with Dublin City University policy on Conflicts of Interest.

I and my co-investigators and/or supporting staff have the appropriate qualifications, experience and facilities to conduct the project set out in the attached application and to deal with any emergencies and contingencies related to the project that may arise. Supervisor(s) signature(s) are required as evidence that they have read and approve the submission.

Please note:

1. *Any amendments to the original approved proposal must receive prior School Ethics Committee approval.*

2. *As a condition of approval investigators are required to document and report immediately to the School of Computing Ethics Committee any adverse events, any issues which might negatively impact on the conduct of the project and/or any complaint from a participant relating to their participation in the study*

Electronic Signature(s):

Principal investigator / Supervisor: ___Alessandra Mileo_____

*Print Name(s) here:*___Alessandra Mileo_____

Date: ___26th April 2023_____

I/We, the students on this proposal, have read and approve this submission

*Student(s) signature(s):*__Adam Tegart_____

*Print Name(s) here:*___Adam Tegart_____

Date: ___26th Apr 2023___

2.   PROJECT OUTLINE

2.1   LAY DESCRIPTION, AIMS & JUSTIFICATION, METHODOLOGY *(up to.100 words)*
Please outline, in terms that any non-expert would understand, what your project is about, including what participants will be required to do. Please explain any technical terms or discipline-specific phrases. State the aims and significance of the project.

> The aim of this project is to understand the extent to which automated methods can be used to find potential concepts for a complex task such as detecting mitotic figures. In this context, a concept is a visual features such as the size, shape or colour of the nucleus within a cell. The participants will assess groups of images that were derived from the publicly available Mitosis Domain Generalization challenge dataset (MIDOG) that were found to be influential and determine if they relate to any visual concepts that are typically used to identify a mitotic figure by experts.

2.2   PARTICIPANT PROFILE
List and very briefly describe each participant group where applicable. For instance, participant group 1 will consist of…, participant group 2 will consist of… etc. Indicate if minors (Under 18) are involved *Provide the number, age range and source of participants. Please provide a justification of your proposed sample size.*

> There will be one set of participants, pathologists/experts. The motivation behind this is that they are professionals who encounter slides containing mitotic figures regularly and they have received the appropriate training and education to identify them. The sample size will be relatively small, consisting of one or two individuals. The source of the participants will be individuals within and/or partnered with the company I carried out my INTRA placement in, Deciphex.

2.3   PARTICIPANT RECRUITMENT
Please provide specific details as to how you will be recruiting participants. How will people be informed that you are doing this research? How will they be approached and asked if they are willing to participate? If you are mailing or phoning people, please explain how you have obtained their names and contact details. If a recruitment advertisement (e.g. through social media, if so include the text at the end of the form) is to be used, please ensure you attach a copy to this application (Approx. 100 words).

> The participants will be recruited through the company I carried out my INTRA with, Deciphex. I will reach out to my previous manager and they will put me in contact with individuals who have sufficient knowledge to evaluate the results.

2.4I   IS IT LIKELY THAT ANY PARTICIPANTS COULD BE CONSIDERED POTENTIALLY VULNERABLE?
Are some or all participants vulnerable in any way? (e.g by virtue of the group they belong to, people who have undergone traumatic or adverse emotional events, people with diminished cognitive ability, power relations between students and participants etc.)?

YES or NO
NO

If Yes, please state and describe what this vulnerability (or vulnerabilities) is and justify why this research is being done with such participants

N/A

2.5   WILL THE IDENTITY OF THE PARTICIPANTS BE PROTECTED?

YES or NO
YES

N/A

IF YOU ANSWERED YES TO 2.5, PLEASE ANSWER THE FOLLOWING QUESTION:

2.6   HOW WILL THE ANONYMITY OF THE PARTICIPANTS BE RESPECTED?
Please bear in mind that where the sample size is very small, it may be impossible to guarantee anonymity/confidentiality of participant identity. Participants involved in such projects need to be advised of this limitation in the Plain Language Statement/Information Sheet. If you intend to fully anonymize the data, please provide details

N/A

2.7   LEGAL LIMITATIONS TO DATA CONFIDENTIALITY
*Participants need to be made aware that confidentiality of information provided cannot always be guaranteed by researchers and can only be protected within the limitations of the law - i.e., it is possible for data to be subject to subpoena, freedom of information claim or mandated reporting by some professions. **This information should be included in your Plain Language Statement and Informed Consent Form.** Depending on the project proposal and academic discipline, you may need to state additional specific limitations.*

State how and where participants will be informed of these limitations

The participants will be made aware of the legal limitations to data confidentiality in the plain language statement and informed consent form as advised.

2.8(a)   EXPLAIN HOW PARTICIPANTS ARE TO BE RECRUITED
Please provide specific details as to how you will be recruiting participants. How will people be informed that you are doing this research? How will they be approached and asked if they are willing to participate? If you are e-mailing, mailing or phoning people, please explain how you have obtained their names and contact details. If a recruitment advertisement is to be used, please ensure you attach a copy to this application.

The participants will be recruited through the company I carried out my INTRA with, Deciphex. I will reach out to my previous manager and they will put me in contact with individuals who have sufficient knowledge to evaluate the results.

2.8(b)   CHILD PARTICIPANTS (anyone under 18 years old)
*If your participants include children, you **must** confirm that you are in compliance with the research specific guidelines as detailed in "Keeping Children Safe - Policies and Procedures supporting Child Protection at DCU" - available at:* ***https://www4.dcu.ie/sites/default/files/policy/157%20-%20child_protection_handbook_rev1%282%29%281%29.pdf***

Please indicate your compliance with the following guidelines:	**Mark here**
We confirm that we have read and agree to act in accordance with the DCU Child Protection policy and procedures	N/A
We confirm that we have put in place safeguards for the children participating in the project	N/A
We confirm that we have supports in place for children who may disclose current or historical abuse (whether or not this is the focus of the project)	N/A

2.9   PLEASE EXPLAIN WHEN, HOW, WHERE, AND TO WHOM RESULTS WILL BE DISSEMINATED, INCLUDING WHETHER PARTICIPANTS WILL BE PROVIDED WITH ANY INFORMATION AS TO THE FINDINGS OR OUTCOMES OF THE PROJECT?

The results will be disseminated at the closing of the project when a paper is submitted around May 9th, this paper may be published dependent on findings. Due to the nature of the questionnaire the participants will know the result of their contribution as they are assessing results on my behalf.

2.10   ARE OTHER APPROVALS REQUIRED TO GAIN ACCESS TO ANOTHER LOCATION, ORGANISATION, SCHOOL ETC.?

YES or NO
NO

If YES, please specify from whom and attach a copy of the approval documentation. If this is not yet available, please explain when this will be obtained.

N/A

3.   RISK AND RISK MANAGEMENT

3.1   EXPLAIN AND JUSTIFY THE STATED LEVEL OF RISK TO PARTICIPANTS
You must provide a justification that the stated level of risk and its corresponding level of review is notification only and not Full Committee or Expedited, as indicated on the cover page of your application. No project is completely without risk. Note that the level of risk may be influenced by the vulnerability of the research group, the methods employed and the nature of the project itself. For further information on risk levels, please refer to the Levels of Review information on the website: https://www.dcu.ie/researchsupport/researchethics.shtml

The level of risk is very low as the participant group is not vulnerable and their expertise will be used in a professional capacity. There will be no questions relating to personal details and the information gathered will be used to advise on classifying the results. This means that no data that would allow the reidentification of participants will be collected.

3.2   POTENTIAL RISKS TO PARTICIPANTS AND RISK MANAGEMENT PROCEDURES
Identify, as far as possible, all potential risks to participants (physical, psychological, social, legal, economic, etc.), associated with the proposed project. Will your project involve deception, investigation of participants involved in illegal activities, performance of any acts which might diminish the self-esteem of participants or cause them to experience embarrassment, regret or depression? Please explain what risk management procedures will be put in place to minimise these risks.

There would typically be a risk of identification due to the small group of participants, so to mitigate this I will not be collecting any personal data or data that would allow reidentification of the participants. There is also risk of the individual becoming tired, so I will ensure that only the necessary questions are asked to keep the length of the questionnaire to a minimum.

3.3   ARE THERE LIKELY TO BE ANY BENEFITS (DIRECT OR INDIRECT) TO PARTICIPANTS FROM THIS RESEARCH?

YES or NO
YES

If YES, provide details

The findings from this project will help to understand the extent to which these methods can be applied to problems in the health space. Additionally, the methods applied are being used in a novel use case and understanding the extent to which they worked well will lead to insight that can be used to direct further research efforts in the area of explainable AI.

3.4 ARE THERE ANY SPECIFIC RISKS TO YOURSELVES IN CARRYING OUT THIS PROJECT?
Examples include use of dangerous materials, asking certain types of questions,The project being undertaken in certain locations, researchers working alone in isolated areas, etc.

YES or NO
NO

If YES, please describe and explain what risk management procedures will be put in place to minimise these risks

N/A

3.5 DEALING WITH ADVERSE/UNEXPECTED OUTCOMES
Please describe what measures/protocols you have put in place in the event that there are any unexpected outcomes or adverse effects to participants arising from involvement in the project.

We agree to regularly meet with our supervisor to monitor the project and enable them to help deal with unexpected outcomes, and this will provide support for participants and monitor the project

YES or NO
YES

3.6 SUPPORT FOR PARTICIPANTS
Depending on risks to participants you may need to consider having additional support for participants during/after the study. Consider whether your project would require additional support, e.g., external counselling available to participants. Please advise what support will be available.

As the participants will be taking part in this questionnaire in a professional capacity there is little risk to the participants. I will liaise with my contact in Deciphex to ensure that the participants have sufficient supports in place internally.

3.7 HOW WILL THE CONDUCT OF THE PROJECT BE MONITORED?
Please explain how the principal investigator will monitor the conduct of the project (especially where several people are involved in recruiting or interviewing, administering procedures, etc.) to ensure that it conforms with the procedures set out in this application. In the case of student projects please give details of how the supervisor(s) will monitor the conduct of the project.

The principal investigator will review the material for the questionnaire to be administered to the participants and be involved in communication with my contact in Deciphex and participants.

3.8 DO YOU PROPOSE TO OFFER PAYMENTS OR INCENTIVES TO PARTICIPANTS?

YES or NO

NO

N/A

3.9 **DO ANY OF THE RESEARCHERS ON THIS PROJECT HAVE A PERSONAL, PHILOSOPHICAL, FINANCIAL, POLITICAL, IDEOLOGICAL, OR COMMERCIAL INTEREST IN ITS OUTCOME THAT MIGHT INFLUENCE THE INTEGRITY OF THE PROJECT OR BIAS THE CONDUCT OR REPORTING OF THE PROJECT, OR UNDULY DELAY OR OTHERWISE AFFECT THEIR PUBLICATION?**

YES or NO
NO

If YES, please specify how this conflict of interest will be addressed

N/A

4. PERSONAL DATA - COMPLIANCE WITH THE GENERAL DATA PROTECTION REGULATION (GDPR)

Applicant declaration:

0	I understand that the proposed project, as set out in this form, is to be carried out by me in my capacity as a student of Dublin City University.	YES or NO	YES

Definition of Personal Data

Personal data is any information about a living person, where that person is either identified or could be identified, from the data itself or when it is combined with other data. Typical examples of personal data in a research context are:

a) paper based records e.g. consent forms, research participant files, patient records, interview notes etc.

b) electronic records e.g. database of participant details, online survey returns, photos, audio & visual recordings, IP addresses, diagnostic / clinical imaging etc.

c) other e.g. genetic data, biometric data, clinical or medical samples etc.

Note: If personal data is to be obtained and / or processed in the course of the proposed research then there are certain legal obligations and principles to be followed. These are set out in the EU 2016 General Data Protection Regulation (GDPR) and associated Irish Law.

Any data that is <u>fully and completely anonymous</u> is not considered to be 'personal data'. However, any data that is merely pseudo-anonymised is deemed to be 'personal data'.

Further information on data protection issues is available from the University's Data Protection Unit (DPU). You should also consider consulting with your Unit's GDPR Advocate for help and advice on filling out this section of the form.

4.1   ASSESSING DATA PROTECTION RISKS & REQUIREMENTS

(A) Your knowledge of Data Protection		
Have you taken and completed the online data protection training course ('Data Protection Course') that is available to all staff and students through the DCU Loop System?	YES or NO	YES

If you answered 'No' to the previous question then the DPU strongly recommends that all applicants complete the course on Loop before completing section # 4 of the REC Application Form.

If you experience difficulties in accessing the Loop course at the link above, please contact the Teaching Enhancement Unit for assistance.

(B) Initial Assessment of whether any of the data to be used in the proposed research is 'Personal Data' (see definition above)				
1	**Will the proposed research include living human subjects?** *Rationale – personal data applies only to living individuals.*		**YES or NO**	**YES**
2	**Will the proposed research use any data that can be linked to an identified, or an identifiable, person?** *Rationale – to be personal data it must be possible to associate it with an identified, or an identifiable, living person.*		**YES or NO**	**NO**
3	**Will the proposed research use any data identifiers that can be linked to a living person? Examples are a participant's name, code or ID number, their address, their IP address etc.** *Rationale: fully anonymised data is not deemed to be 'personal data' but data that has been deemed to be merely pseudo-anonymised is deemed to be 'personal data'.*		**YES or NO**	**NO**

If you answered 'Yes' to any of the questions 1 to 3 in sub-section (B), then continue to sub-section (C) and answer questions 1-8. If you answered 'No' to all of the questions 1 to 3 in sub-section (B), then proceed directly to section # 5 of this Application Form.

(C) Assessing the degree of risk inherent in the personal data				
1	**Will the proposed research involve the use of <u>personal data</u> on individuals that reveals any of the following attributes or characteristics about them?** **(State 'Yes' or 'No' as appropriate to all of the following)**			**N/A**
		Racial or Ethnic Origin	**YES or NO**	**NO**

			YES or NO	NO
		Political Opinions	YES or NO	**NO**
		Religious or Philosophical Beliefs	YES or NO	**NO**
		Trade Union Membership	YES or NO	**NO**
		Genetic Data	YES or NO	**NO**
		Biometric Data	YES or NO	**NO**
		Data Concerning Health	YES or NO	**NO**
		Data concerning a Person's Sex Life or Sexual Orientation	YES or NO	**NO**
2	**Will the proposed research involve the use of <u>personal data</u> relating to children or vulnerable individuals?** *A child, for data protection purposes, is defined as an individual below 18 years of age. Where the processing relates to 'electronic marketing' the age limit is reduced to 16 years. A vulnerable individual may be anyone who is unable to consent to, or to oppose, the processing of his or her data for any reason, including disability.*		YES or NO	**NO**
3	**Will the proposed research involve the use of data relating to an individual's criminal convictions and / or offences?**		YES or NO	**NO**

4	**Will the proposed research involve the large-scale processing of <u>personal data</u>?** *This may include: a wide range or large volume of personal data; processing which takes place over a large geographical area; processing where a large number of people are affected (e.g. over 100 individuals); or where the processing is extensive or it has potential long-lasting effects on individuals.*	**YES or NO**	**NO**
5	**Will the proposed research involve any form of <u>automated processing</u> of personal data?** *In particular, to analyse or predict aspects concerning that person's performance at work, economic situation, health, personal preferences, interests, reliability, behaviour, location or movements.*	**YES or NO**	**NO**
6	**Will the proposed research involve the sharing or transferring of any personal data to a 3rd party outside of DCU?** *For example, other research partners, providers of translation or transcription services, etc.* *For clarity, this question is not intended to refer to any standard software services already provided by DCU, for example the university's email system or its cloud-based storage provider (Google Drive).*	**YES or NO**	**NO**
7	**Will the proposed research require the sharing or processing of personal data outside the EU or the EEA? (e.g. the US, the UK, Canada, Australia, China etc.)** *The EEA refers to the 'European Economic Area' (i.e. the EU plus Norway, Liechtenstein and Iceland).*	**YES or NO**	**NO**

8	**Will the proposed research involve the matching or combining of separate datasets of information on individuals in a way that would exceed their reasonable expectations of privacy?** *This is especially important where two or more previously anonymous datasets are combined in such a way so as to allow for the identification of individuals. An example would be combining mobile phone location data along with any other dataset to identify individuals.*	**YES or NO**	**NO**

Important Point: Next Step

If you answered 'Yes' to one or more of the questions 1 to 8 in sub-section (C) **You should consult with your Supervisor / Principal Investigator to who** will assess whether there are any further data protection issues to be addressed or additional procedures to be followed.

Note 1: What does 'Minor' and 'Vulnerable Individual' mean?

A **minor** is defined as an individual below 18 years of age. Where the processing relates to 'electronic marketing' the age limit is reduced to 16 years. A **vulnerable individual** may be anyone who is unable to consent to, or oppose, the processing of his or her personal data for any reason. Both of these are of particular importance if the project compels the provision of data from individuals.

Note 2: What does 'large scale processing' mean?

The GDPR does not define what constitutes large-scale. EU guidance recommends that the following factors, in particular, be considered when determining whether the processing is carried out on a large scale:
- the number of data subjects (either as a specific number or proportion of the relevant population);
- the volume of data and/or the range of different data items being processed;
- the duration, or permanence, of the data processing activity; &
- the geographical extent of the processing activity.

Examples of large-scale processing include, but are not limited to:
- processing of patient data in the regular course of business by a hospital;
- processing of travel data of individuals using a public transport system (e.g. tracking via travel cards);
- processing of real time geo-location data of customers of an international fast food chain for statistical purposes by a processor specialised in these activities;
- processing of customer data in the regular course of business by an insurance company or a bank;
- processing of personal data for behavioural advertising by a search engine; &
- processing of data (content, traffic, location) by telephone or internet service providers.

Examples that do **not** constitute large-scale processing include, but are not limited to:
- processing of patient data by an individual physician; and
- processing of personal data relating to criminal convictions and offences by an individual lawyer.

	B. Applicant Data Protection Assessment Questionnaire – Part II		
5(a)	Does your project include the use of Personal Data of individuals which reveals any of the attributes or characteristics below? If 'Yes,' please indicate which will be used in your project **(tick all that apply):**	**YES or NO**	**NO**

	racial or ethnic origin	YES or NO	NO
	political opinions	YES or NO	NO
	religious or philosophical beliefs	YES or NO	NO
	trade union membership	YES or NO	NO
	genetic data	YES or NO	NO
	biometric data	YES or NO	NO
	data concerning health	YES or NO	NO
	data concerning a natural person's sex life or sexual orientation	YES or NO	NO
5(b)	Does your project include the use of Personal Data relating to minors or vulnerable individuals? *(See **Note 1**, below)*	YES or NO	NO
6	Does your project include the use of Personal Data of individuals relating to their criminal convictions and/or offences?	YES or NO	NO
7	Does your project include large-scale processing of personal data relating to living individuals? *This may include: a wide range or large volume of personal data; processing which takes place over a large geographical area; or where a large number of people are affected (e.g. over 100 individuals); or where the processing is extensive or has long-lasting effects. (See **Note 2**, below)*	YES or NO	NO
8	Does your project include any form of automated processing of personal data, used to evaluate certain personal aspects relating to a living individual? *In particular, to analyse or predict aspects concerning that person's performance at work, economic situation, health, personal preferences, interests, reliability, behaviour, location or movements*	YES or NO	NO
9	Does your project include any partners which are third parties outside of DCU? *e.g. Research partners, third party software providers or other providers such as translation or transcription services, etc.*	YES or NO	NO
10 (a)	Does your project involve the sharing or processing of Personal Data outside the EU or the EEA? *i.e. the EEA is the European Economic Area (the EU plus Norway, Liechtenstein and Iceland)*	YES or NO	NO

10 (b)	If 'Yes', please state which non-EU or EEA country is involved:	N/A	
11	Does the project require the matching or combining of separate datasets of information on individuals in a way that would exceed their reasonable expectations of privacy? *An example would be combining mobile phone location data along with any other dataset to identify individuals.*	**YES or NO**	**NO**

If you answered 'Yes' to one or more of these questions, you should make sure that you have strong and secure data privacy risk mitigation safeguards in place, discuss these with your supervisor.

4.2   WILL ANONYMISATION OR PSEUDONYMISATION OF THE PERSONAL DATA, WHERE APPLICABLE, BE UNDERTAKEN?
Anonymisation is the process of removing personal identifiers, both direct and indirect, that may lead to an individual being identified. Pseudonymisation is the processing of personal data in such a manner that the personal data can no longer be attributed to a specific living individual without the use of additional information, provided that such additional information is kept separately and is subject to technical and organisational measures to ensure its security.

YES or NO
NO

If YES, please explain below the methods by which you intend to anonymise/pseudonymise the personal data:

N/A

5.   DATA/SAMPLE STORAGE, SECURITY AND DISPOSAL

For the purpose of this section the term 'Data' includes personal data that is in a raw or a processed state (e.g. interview audiotape, transcript or analysis, etc.). The term 'Samples' include body fluids and/or tissue samples.

5.1   HOW AND WHERE WILL THE DATA/SAMPLES BE STORED?
DCU recommends that any data stored electronically offsite should utilise the DCU Google Drive. Alternative offsite storage will need to be justified and must meet data protection and GDPR compliance requirements.

The data will be stored on the DCU Google Drive and shared with the participants as needed.

5.2   WHO WILL HAVE ACCESS TO DATA/SAMPLES?
If people other than the main researchers have access, please name who they are and explain for what purpose.

Only the main researcher and principal investigator will have access to the data.

5.3   HOW LONG IS THE DATA TO BE HELD OR RETAINED?
Note that, with very few exceptions, Personal Data may not be retained indefinitely. It is up to the project team to establish an upper retention limit for each category of Personal Data used within the project and to ensure it is applied at the expiry of that limit. The School of Computing Research Ethics Committee recommends that Personal Data is retained until after the Progression and Awards Board for the current academic year.

The data is not personal in nature as the participants are answering the questionnaire in a professional capacity. The data is anonymised at the source (as participants don't supply personal data) and no personal identifiable data, direct or indirect, is stored at any point in time.

5.4 **IF YOUR PROJECT DOES INVOLVE THE USE OF PERSONAL DATA THEN WILL THIS BE USED AT A LATER DATE FOR THE PURPOSE OF PUBLICATION OF THE RESULTS OF THE PROJECT?**

YES or NO
N/A

Where it is intended that the personal data used in the project will be used at a later date for the purposes of publication please explain how consent to do so will be obtained.

N/A

5.5 **IF THE DATA/SAMPLES ARE TO BE DISPOSED OF AT THE END OF THE PROJECT PLEASE EXPLAIN HOW, WHEN AND BY WHOM THIS WILL BE DONE?**

*Note that simply deleting files is not sufficiently secure. The additional steps to be taken to maintain data security should be given. **Personal data** must be disposed of in a safe and secure manner at the end of its retention period. If the data is stored in (a) a paper-based format, then shredding or disposal via a secure bin is recommended; or (b) in an electronic-based format, then deletion of the record or the full anonymization of the data is recommended. If data/samples are **not** being disposed of, please justify that intention.*

How will the data/samples be disposed of? Please describe the means by which the personal data will be deleted or destroyed. This includes personal data held in hard copy and digital formats.	As mentioned, no personal data or data that could lead to the reidentification of individuals is being collected through this questionnaire. So this justifies the choice not to dispose of the data collected from this questionnaire.
By whom will the data/samples be disposed? Please indicate the designated team member(s) with responsibility for deletion and/or destruction of the research project's personal data.	N/A

6. PLAIN LANGUAGE STATEMENT *(Attach to this document. Up to a max of 400 words)*

A Plain Language Statement (PLS) should be used in all cases. This is written information in plain language that you will be providing to participants, outlining the nature of their involvement in the project and inviting their participation. The PLS should specifically describe what will be expected of participants, the risks and inconveniences for them, and other information relevant to their involvement. Please note that the language used must reflect the participant age group and corresponding comprehension level– if your participants have different comprehension levels (e.g. both adults and children) then separate forms should be prepared for each group. The PLS can be embedded in an email to which an online survey is attached, or handed/sent to individuals in advance of their consent being sought. See link to sample templates on the website: https://www.dcu.ie/researchsupport/ethicsapproval.shtml

PLEASE CONFIRM WHETHER THE FOLLOWING ISSUES HAVE BEEN ADDRESSED IN YOUR PLAIN LANGUAGE STATEMENT/ INFORMATION SHEET FOR PARTICIPANTS:
Note that this list is a check-list of all of the things that you should include in your plain language statement, if they are relevant (they are in most cases). In the earlier sections of this form you have already written the text that can be used to create your plain language statement. References to the relevant sections are provided on each line.

	YES or NO
Introductory Statement (Student(s) and supervisor names, school, title of the project) [Table, p 1]	**YES**
What is this project about? [section 2.1]	**YES**
Why is this project being conducted? [section 2.1]	**YES**
What will the participant be expected to do/have to do if they decide to participate in the study?[section 2.1]	**YES**
How will their privacy be protected? [section 2.5, section 2.6]	**YES**
How will the data be used and subsequently disposed of? [section 5.3]	**YES**
What are the legal limitations to data confidentiality? [section 2.7]	**YES**
Are there any benefits of taking part in the study? [section 3.3]	**YES**
Are there any risks of taking part in the study? [section 3.2]	**YES**
Confirmation that participants can change their mind at any stage and withdraw from the study [see plain language statement template, appendix 1]	**YES**
How will participants find out what happens with the project? [section 2.9]	**YES**
Contact details for further information [see plain language statement template, appendix 1]	**YES**

If any of these issues are marked NO, please justify their exclusion:

N/A

7. INFORMED CONSENT FORM *(Attach to this document. Approx. 300 words, see appendices 2 and 3 for templates.)*

In most cases where interviews or focus groups are taking place, an Informed Consent Form is required. This is an important document requiring participants to indicate their consent to participate in the study and give their signature. In cases where an anonymous questionnaire is being used, it is not enough to include a tick box in the questionnaire. Participants should indicate their consent to each aspect of the research in a staged manner by checking mandatory checkboxes.
See link to sample templates on the website: https://www.dcu.ie/researchsupport/ethicsapproval.shtml

NB – IF AN INFORMED CONSENT FORM IS NOT BEING USED, THE REASON FOR THIS MUST BE JUSTIFIED HERE.

N/A

8.   ASSENT FORM & PLAIN LANGUAGE STATEMENT FOR CHILDREN *(Attach to this document.)*

A child specific Plain Language Statement (PLS) should be used in project where children will be involved. The PLS must be written in a way that is understandable for children within your targeted age group. It also must state, in plain language, the nature of their involvement in the project and inviting their participation. The PLS should specifically describe what will be expected of participants, the risks and inconveniences for them, and other information relevant to their involvement. In addition, child participants should also be provided with an Assent Form. Parents/guardians will be provided with the Informed Consent Form, but each child should provide assent before taking part in the project. The Assent Form needs to be understandable to the age-group you are targeting. See link to sample templates on the website: https://www.dcu.ie/researchsupport/researchethics.shtml

NB – IF AN ASSENT FORM IS NOT BEING USED, THE REASON FOR THIS MUST BE JUSTIFIED HERE.

N/A

9.   SUBMISSION CHECKLIST *(Attach to this document)*

Please confirm that all supplementary information is included in your application (in electronic copy). If questionnaire or interview questions are submitted in draft form, please indicate this by putting (draft) after YES. A copy of the final documentation must be submitted for final approval when available.

My application has been collated as one electronic PDF file which includes the following documentation:	INCLUDED (mark as YES)	NOT APPLICABLE (mark as N/A)
Recruitment advertisement [consistent with section 2.3]		N/A
Plain language statement/Information Statement [see section 6 and appendix 1]	**YES**	
Informed Consent form [see appendices 2 and 3]	**YES**	
Informed Assent form (children only)		N/A
Evidence of external approvals related to the research [see sections 1.1 and 2.10]		N/A
Questionnaire/Survey	**YES**	
Interview/Focus Group Questions		**N/A**

Appendix 1
DUBLIN CITY UNIVERSITY
Sample Template – Plain Language Statement (Up to a max of 400 words)

A Plain Language Statement (PLS) should use language that reflects the participant age group and corresponding comprehension level. It should contain the following information. The headings are there for guidance and do not need to be included in your form.

Introduction to the Study
Identify the Study Title, the university department involved, the student(s) and supervisor

Data Protection/Privacy Notice (Personal Data – GDPR Compliance)
An appropriate Privacy Notice is the means by which data subjects are informed about the use of their data. If personal data is being collected and processed, please refer to https://www.dcu.ie/ocoo/dp/guides.shtml for advice and include the following information in the PLS:

- *The identity of the Data Controller/Joint Data Controller and Data Processor should be clearly. stated. The Data Controller will always be DCU (where the researcher is a DCU researcher), the PLS should identify this and also the name of the project, team and School/Unit. A data processor may hold or process personal data but does not exercise responsibility for or control over the personal data, for example, a transcription service, or a software or cloud hosting company. A Data Processor cannot be an employee of the Data Controller.*
- *The identity of the DCU Data Protection Officer – Mr. Martin Ward (data.protection@dcu.ie Ph: 7005118 / 7008257)*
- *The purpose of the data processing i.e. the reasons why the data is being requested and the purpose to which it will be applied.*
- *The reason(s) for which the data will be processed or held*
- *The categories or types of personal data to be processed*
- *The details of any third parties (i.e. data processors) with whom the data will be shared or transferred, and the reasons for sharing*
- *The details of any external (i.e. non-DCU) parties with whom the data will be shared or transferred, and the reasons for sharing*
- *Where relevant, details of any intention to transfer the data to other countries, especially if outside of the EEA (European Economic Area), and the basis for such transfers*
- *The retention period, or the criteria used to determine retention periods*
- *The right of the individual to lodge a complaint with the Irish Data Protection Commission*
- *Information on the rights of the data subject - Individuals' have the right to access their own personal data and PLS should inform them how to do this and who to contact (DCU Data Protection Unit).*
- *Information on their rights to withdraw consent and who to contact to withdraw consent. In some cases it may be possible for participants to withdraw their consent to the use of their data*
- *If it is intended that the data be used for future studies, you must specify the general parameters of the future further project uses to which the participant's project data may be put.*
- *In cases where personal data will later be anonymized (e.g. for statistical or aggregated data), it is best practice to describe this, so that the participant is fully informed.*

Advice as to whether or not data is to be destroyed after a minimum period
Define when data will be destroyed after the end of the project

Details of what participant involvement in the Study will require
E.g., involvement in interviews; completion of questionnaire; audio/video-taping of events, and the estimated time commitment for the activities

Potential risks to participants from involvement in the Study (if greater than that encountered in everyday life)

Any benefits (direct or indirect) to participants from involvement in the Study

Advice as to arrangements to be made to protect confidentiality of data, including that confidentiality of information provided is subject to legal limitations
Participants need to be made aware that confidentiality of information provided cannot always be guaranteed by researchers – please include the following statement:

"Confidentiality of information can only be protected within the limitations of the law - i.e., it is possible for data to be subject to subpoena, freedom of information claim or mandated reporting by some professions".
Depending on the project proposal and academic discipline, you may need to state additional specific limitations.

Statement that involvement in the Study is voluntary
State that participants may withdraw from the Study at any point. You should explain to the participant that their participation in the project will end, at the point they withdraw, and refer back to the data protection/privacy notice as to what will happen regarding their data. For example, withdrawing consent may mean that no future data collection will take place but previously collected data will still be processed etc.

Any other relevant information – e.g.
- *if the sample size is small, advice to participants that this may have implications for privacy/anonymity*
- *if participants are in a dependent relationship with any of the researchers, a clear statement that their involvement/non-involvement in the project will not affect their ongoing assessment/grades/management*

A Plain Language Statement must end with the following statement:

If participants have concerns about this study and wish to contact an independent person, please contact:
The Secretary, Dublin City University Research Ethics Committee, c/o Research and Innovation Support, Dublin City University, Dublin 9. Tel 01-7008000, e-mail rec@dcu.ie

Appendix 2
DUBLIN CITY UNIVERSITY

Sample Template – Informed Consent Form (approx. 300 words)

An Informed Consent Form should generally contain the information detailed below. It should be written in the first person, e.g. "I will be asked to attend…I may withdraw from the study at any point…..I am aware that the data…etc." The headings are there for guidance and do not need to be included in your form.

Study Title
Also identify the school/centre involved, the principal investigator and any other investigators.

Clarification of the purpose of the study
If personal data is being collected and processed, please ensure that the participants acknowledge the identity of the data controller and the purposes of the processing for which the personal data are intended

Confirmation of particular requirements as highlighted in the Plain Language Statement
Requirements may include involvement in interviews, completion of questionnaire, audio/video-taping of events etc.. Getting the participant to acknowledge requirements is preferable, e.g.

Participant – please complete the following (Circle Yes or No for each question)	
I have read the Plain Language Statement (or had it read to me)	*Yes/No*
I understand the information provided	*Yes/No*
I understand the information provided in relation to data protection	*Yes/No*
I have had an opportunity to ask questions and discuss this study	*Yes/No*
I have received satisfactory answers to all my questions	*Yes/No*
I am aware that my interview will be audiotaped	*Yes/No*

Confirmation that involvement in the Study is voluntary
E.g.I may withdraw from the Study at any point.

Confirmation of arrangements to be made to protect confidentiality of data, including that confidentiality of information provided is subject to legal limitations

Confirmation of arrangements regarding retention/disposal of data

Confirmations relating to any other relevant information as indicated in the PLS
E.g. I consent to the use of my data for future studies within the following parameters (provide detail)

<u>Signature:</u>

I have read and understood the information in this form. My questions and concerns have been answered by the researchers, and I have a copy of this consent form. Therefore, I consent to take part in this project

Participants Signature: _____

Name in Block Capitals: _____

Witness: _____

Date: _____

Appendix 3
Anonymous Online Consent Form Template

In cases where an anonymous questionnaire is being used, researchers are required to provide a separate tick box for each statement that the participant is being asked to consent to/acknowledge. Each statement must be included as an essential field in order to ensure that full informed consent has been obtained. (see example below).

An Informed Consent Form should generally contain the information detailed below. It should be written in the first person, e.g. "I will be asked to attend…I may withdraw from the study at any point…..I am aware that the data…etc." The headings are there for guidance and do not need to be included in your form.

Study Title
Also identify the school/centre involved, the supervisor and any students.

Clarification of the purpose of the study

Confirmation of particular requirements as highlighted in the Plain Language Statement
Getting the participant to acknowledge requirements is mandatory, Participants should not be able to access the survey until they have agreed to all items and indicated their consent.

Example:

Participant – please complete the following (by clicking Yes/No for each question)

I have read the Plain Language Statement (or had it read to me) *

◉ Yes
○ No

I understand the information provided *

○ Yes
○ No

I have had an opportunity to ask questions and discuss this study *

○ Yes
○ No

I understand the information provided in relation to data protection *

○ Yes
○ No

I have received satisfactory answers to all my questions *

○ Yes
○ No

I understand I may withdraw from the Research Study at any point *

○ Yes
○ No

I have read and understand the arrangements to be made to protect confidentiality of data, including that confidentiality of information provided is subject to legal limitations *

○ Yes
○ No

I have read and understand confirmations relating to any other relevant information as indicated in the PLS *

○ Yes
○ No

I consent to participate in this research study *

○ Yes
○ No

Explainable AI in Pathology - Concept Based Explainability for Mitotic Figure Detection in Whole Slide Images.

<u>Plain Language Statement</u>

This study is being carried out as part of a Data Science final year project within the School of Computing, DCU. The student involved is Adam Tegart (adam.tegart2@mail.dcu.ie) and the supervisor for the project is Dr Alessandra Mileo (alessandra.mileo@dcu.ie).

This project concerns the investigation into how effective explainability methods based on visual concepts/features work for the use case of detecting mitotic figures. These concept-based explainability methods work by creating an understanding of what a visual concept is by using image examples, in this use case images of an abnormal cell membrane for example.

The motivation for carrying out this project is to understand how well this approach can be automated for this complex use case, eliminating the need to gather images that relate to a concept.

The participants will receive a questionnaire containing groups of images (derived from the publicly available Mitosis Domain Generalization challenge dataset) that were found through an automated approach and were found to be influential. They will be asked to determine if these groups of influential images relate to a visual concept that experts would typically use to identify a mitotic figure and select any images that do not fall within this concept.

The privacy of participants will be protected as no personal data or data that would allow for reidentification of the participant will be collected. The participants will be required to answer questions in a professional capacity only.

The collected data will be used as a guide for interpreting the results and understanding how well this automated approach can be applied.

The confidentiality of information provided cannot always be guaranteed by researchers and can only be protected within the limitations of the law - i.e., it is possible for data to be subject to subpoena.

The results of this survey will aid in determining the effectiveness of this method and will help to guide future research efforts in the area of explainable AI.

There are no perceived risks for participating in this study greater than the risks we encounter in everyday life. Participation in this study is voluntary and consent can be withdrawn at any stage, resulting in the collected data being purged.

The findings from this project will be divulged to participants upon request and may additionally be published.

If participants have concerns about this study and wish to contact an independent person,
please contact:
The Secretary, Dublin City University Research Ethics Committee, c/o Research and Innovation Support, Dublin City University, Dublin 9. Tel 01-7008000, e-mail rec@dcu.ie

<u>Informed Consent</u>

School of Computing, DCU. Adam Tegart (Student), Dr Alessandra Mileo (Supervisor).

The purpose of this questionnaire is to gain insight into the results obtained during the course of the project from the perspective of an expert. The results of this questionnaire will help to quantify the efficiency of these methods for finding concepts automatically and ensure that they align with current knowledge within the field. Quantifying and understanding the utility of the methods applied will allow for insights into future research ideas in the field and guide this future research.

I have read the plain language statement (or had it read to me)	☐
I understand the information provided	☐
I have had the opportunity to ask question and discuss the study	☐
I understand the information provided in relation to data protection	☐
I have received satisfactory answers to all my questions	☐
I understand that I have the right to withdraw from the study at any stage	☐
I understand that the confidentiality of the information provided is subject to legal limitations	☐
I consent to participate in this study	☐

The questions in this questionnaire are composed as follows:

- A grid of 30 images (segments of annotations) that have been sampled from a group that is found to be influential are displayed along with the context image (complete image, zoomed out for context) to help see where the segment is in relation to the mitotic figure.
- You will be asked if these images (or most of at least) represent some visual feature of a mitotic figure. Please try to be as descriptive as you can.
- You will then be asked if any of the images don't belong in the grouping, meaning an image does not represent the concept that this group should supposedly be representing.

The images in the grid are indexed from 0-29, and these can be used to reference images that do not belong in a set. There is a black box on the context image that represents the original annotation. The image was segmented at the level of the original annotation and at the context level (zoomed out). The motivation was to capture the different levels of visual features, such internal cell features in the annotation and the position in relation to nearby cells in the context.

The patch segment can be found in the context image by matching it against the annotation (inner) or context box (outer), the example below shows an annotation segment, which can be found at the top right of the inner/annotation box. The image index (1199_005 below) captures the annotation index (1199) and segment index (005). If the segment index is 9 or larger the segment is likely a context one (zoomed out), so is shown in relation to the outer box.
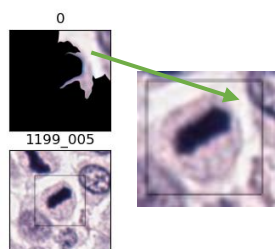
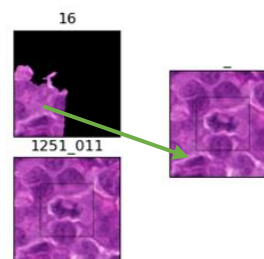Fig1. Example of an annotation segment (within inner box).

Fig2. Example of a context segment (can contain annotation).

This will be carried out for 4 different tissue types, totalling 10 concepts:
- Canine Lymphoma – 5 concepts
- Canine Cutaneous Mast Cell Tumour – 2 concepts
- Canine Lung Cancer – 1 concept
- Human Neuroendocrine Tumour – 2 concepts

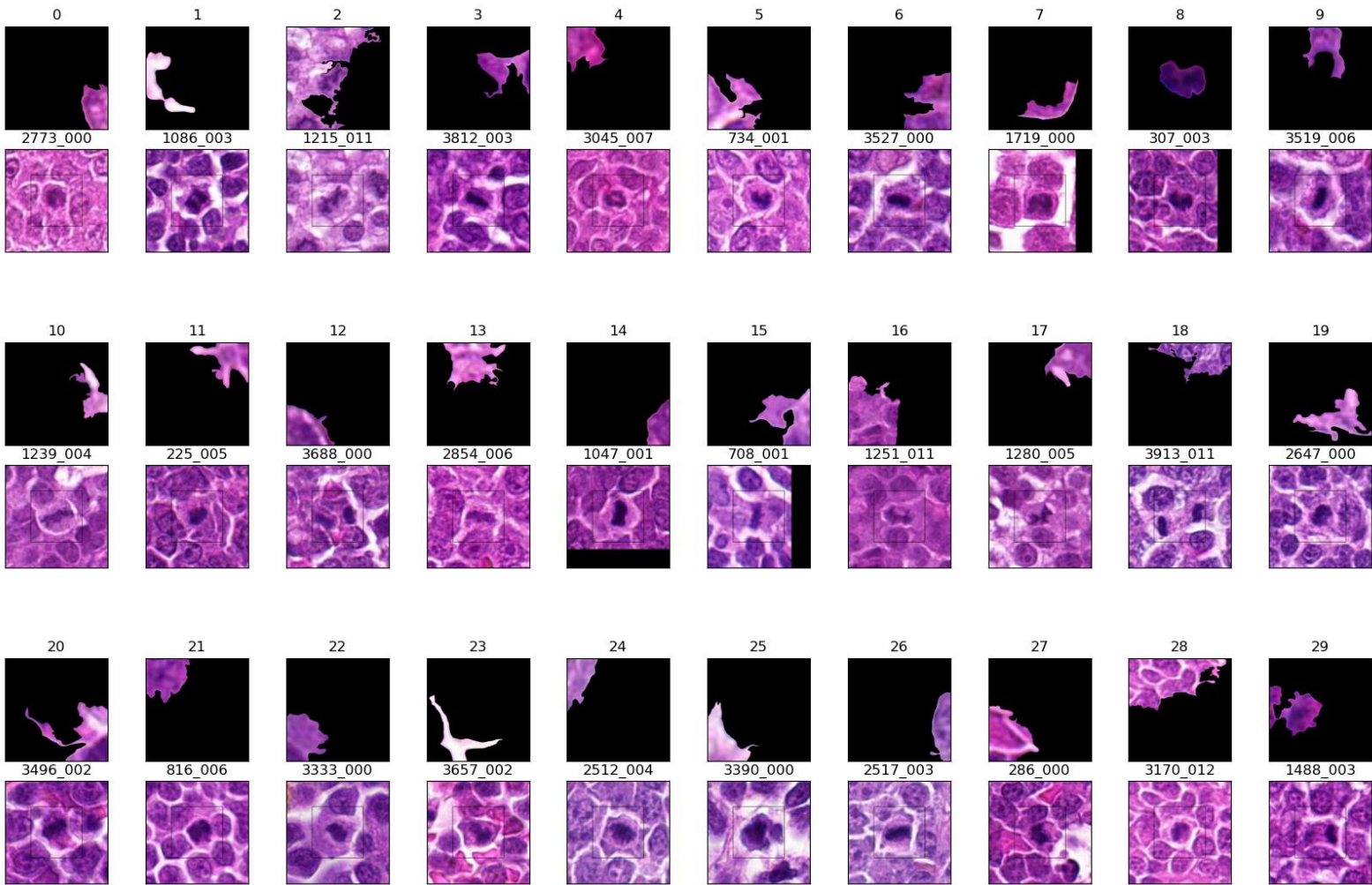Q1. Could you describe the visual features that are typically used to identify a mitotic figure?

Canine Lymphoma

Q2a.

Is there any visual feature that these images (or most of) represent? If not, is there any pattern visible?
Please answer below.

backbone.body.layer4.2.conv1 mitotic figure_concept50

Q2b.

Are there any images in the above group that are not representative of this feature/concept? Please give the grid
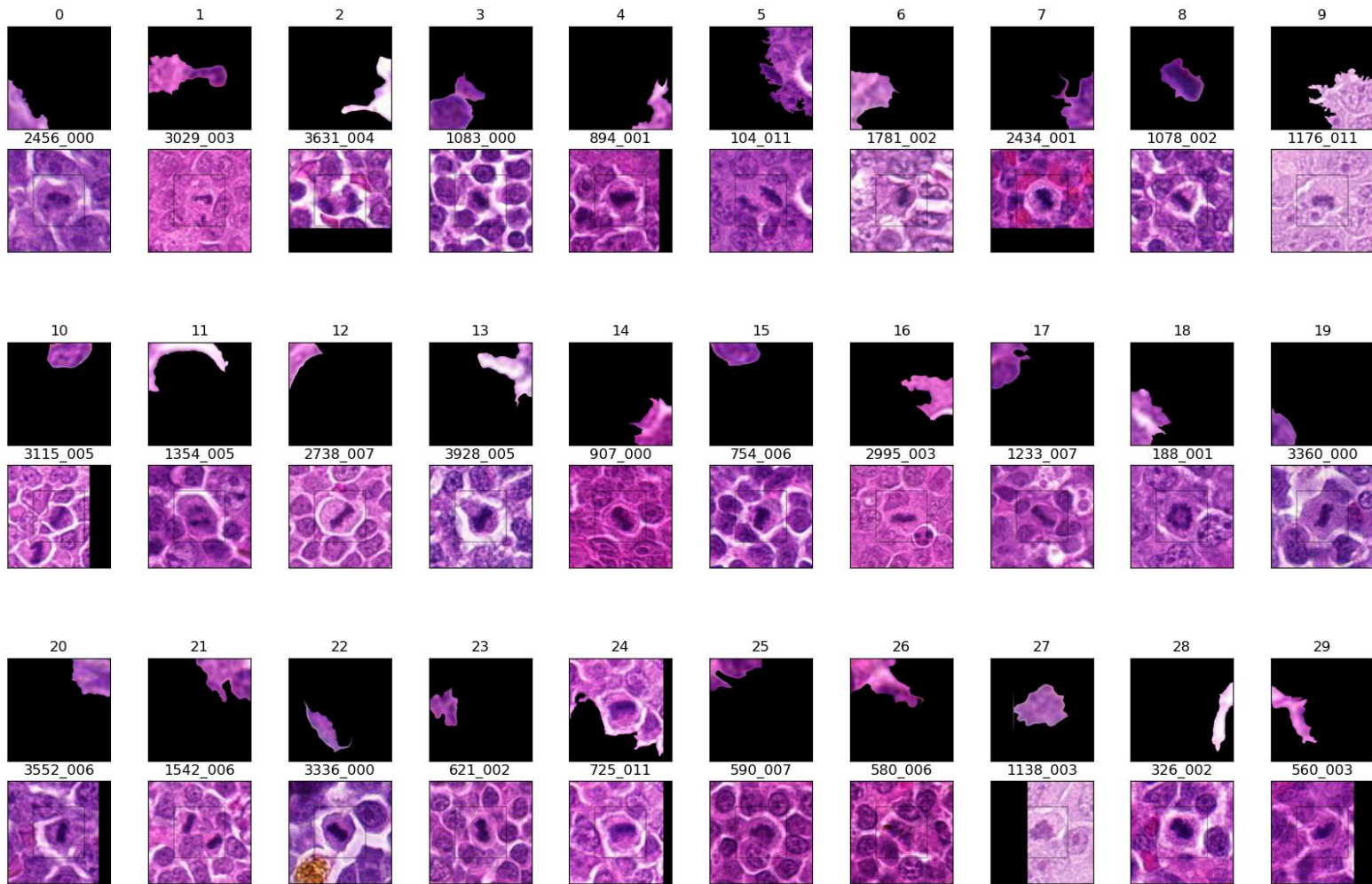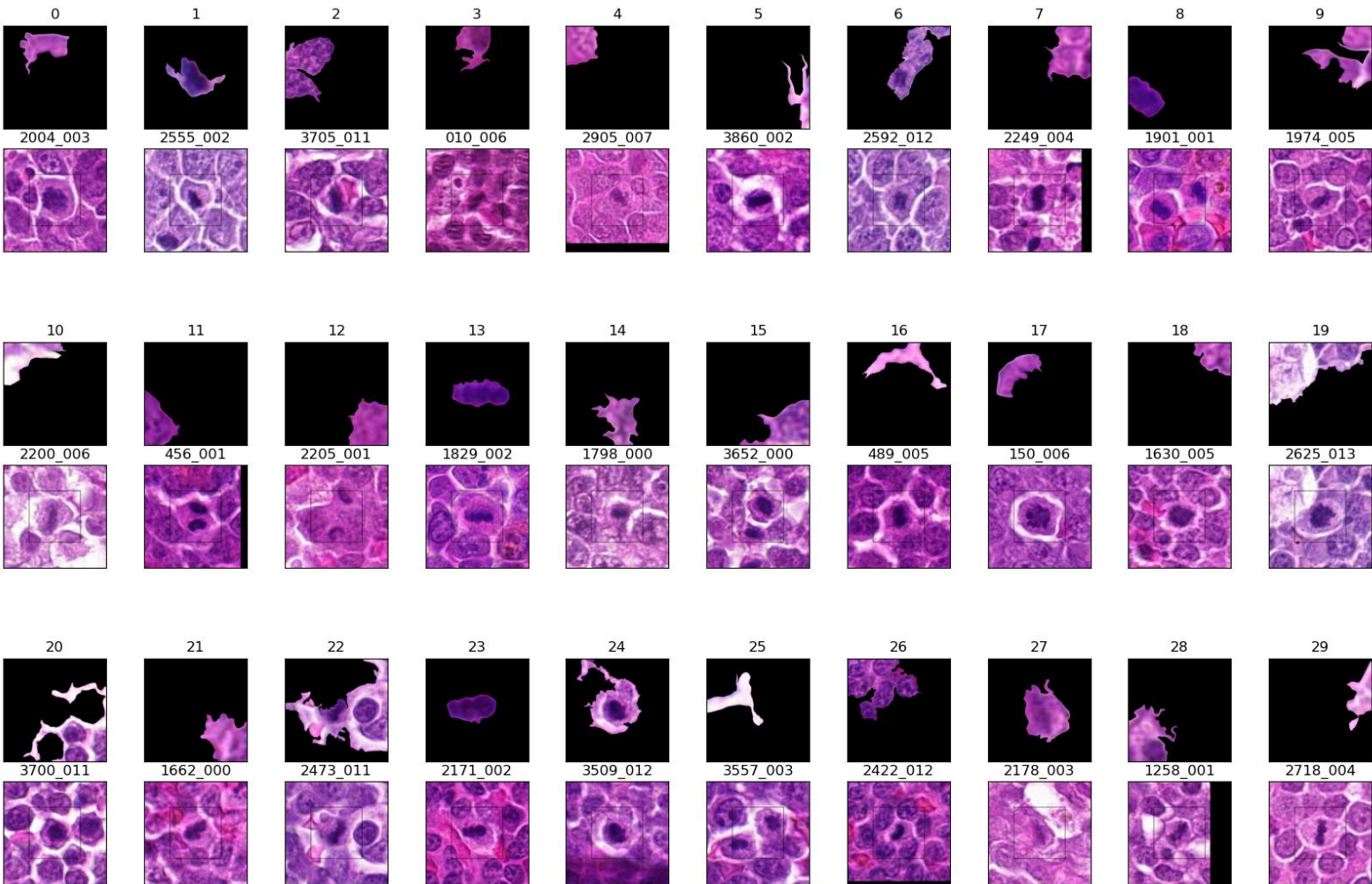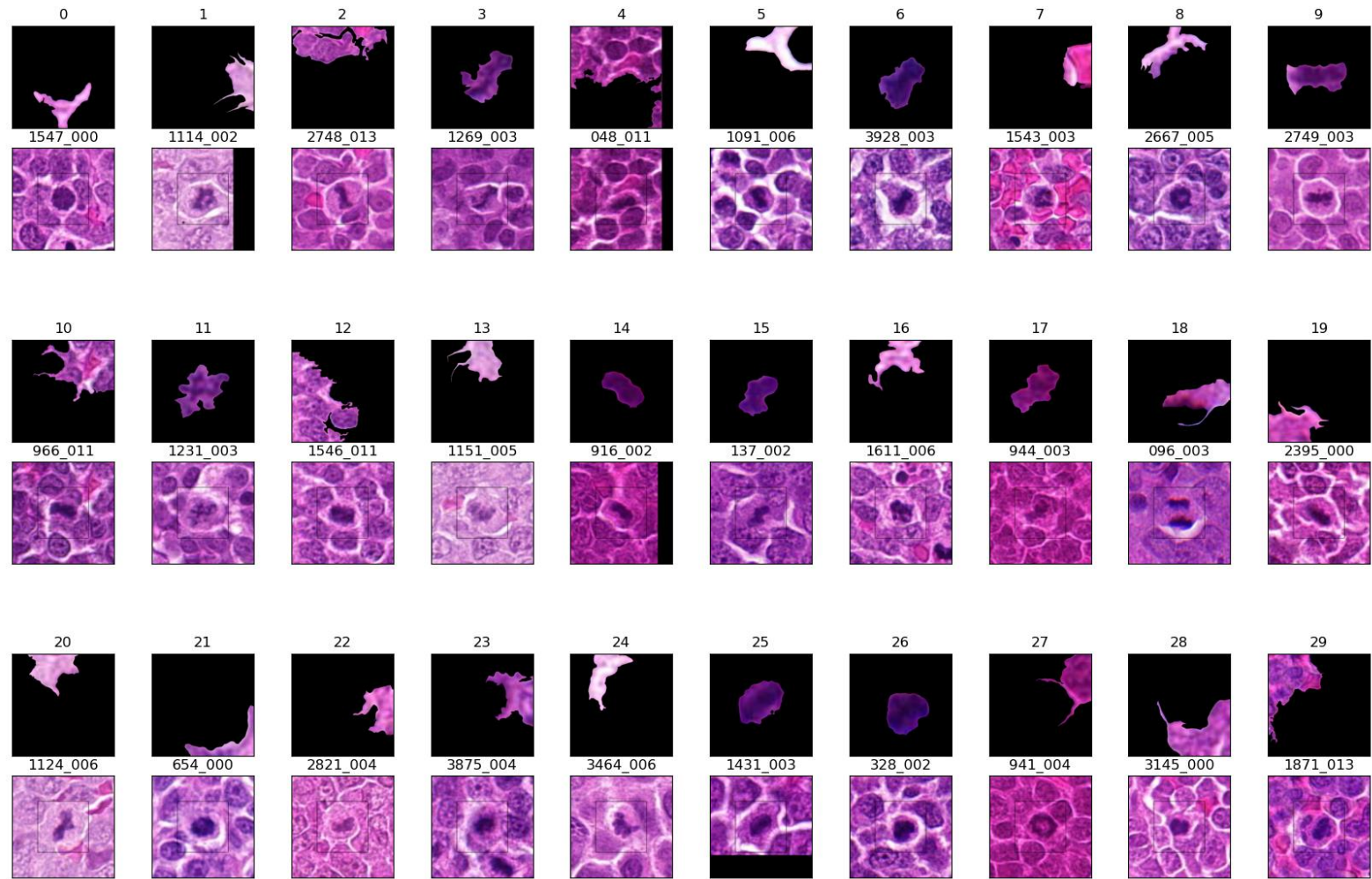index and briefly explain.

Canine Lymphoma

Q3a.

Is there any visual feature that these images (or most of) represent? If not, is there any pattern visible?
Please answer below.

backbone.body.layer4.2.conv1 mitotic figure_concept9

Q3b.

Are there any images in the above group that are not representative of this feature/concept? Please give the grid index and briefly explain.
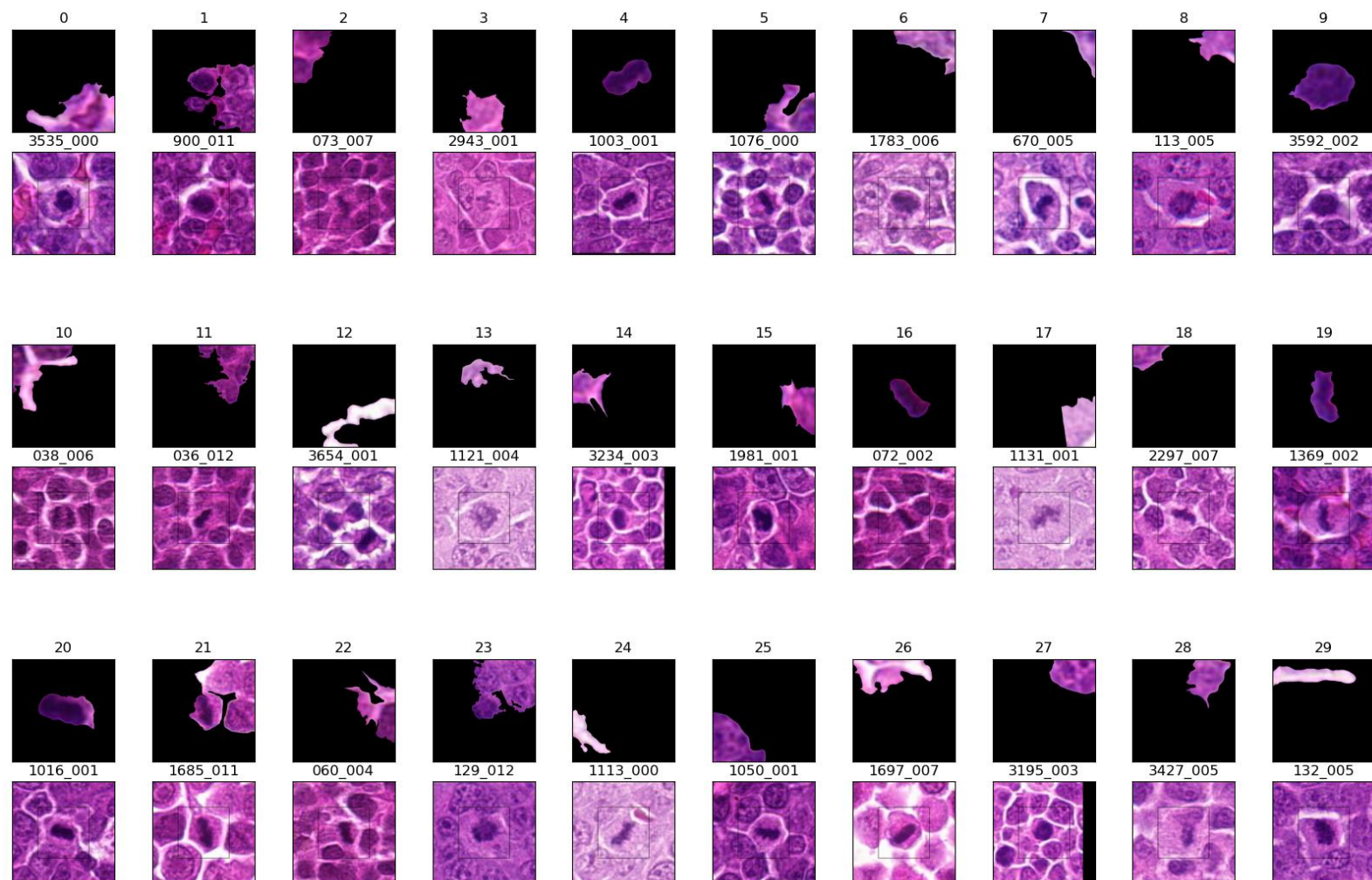
Canine Lymphoma

Q4a.

Is there any visual feature that these images (or most of) represent? If not, is there any pattern visible?
Please answer below.

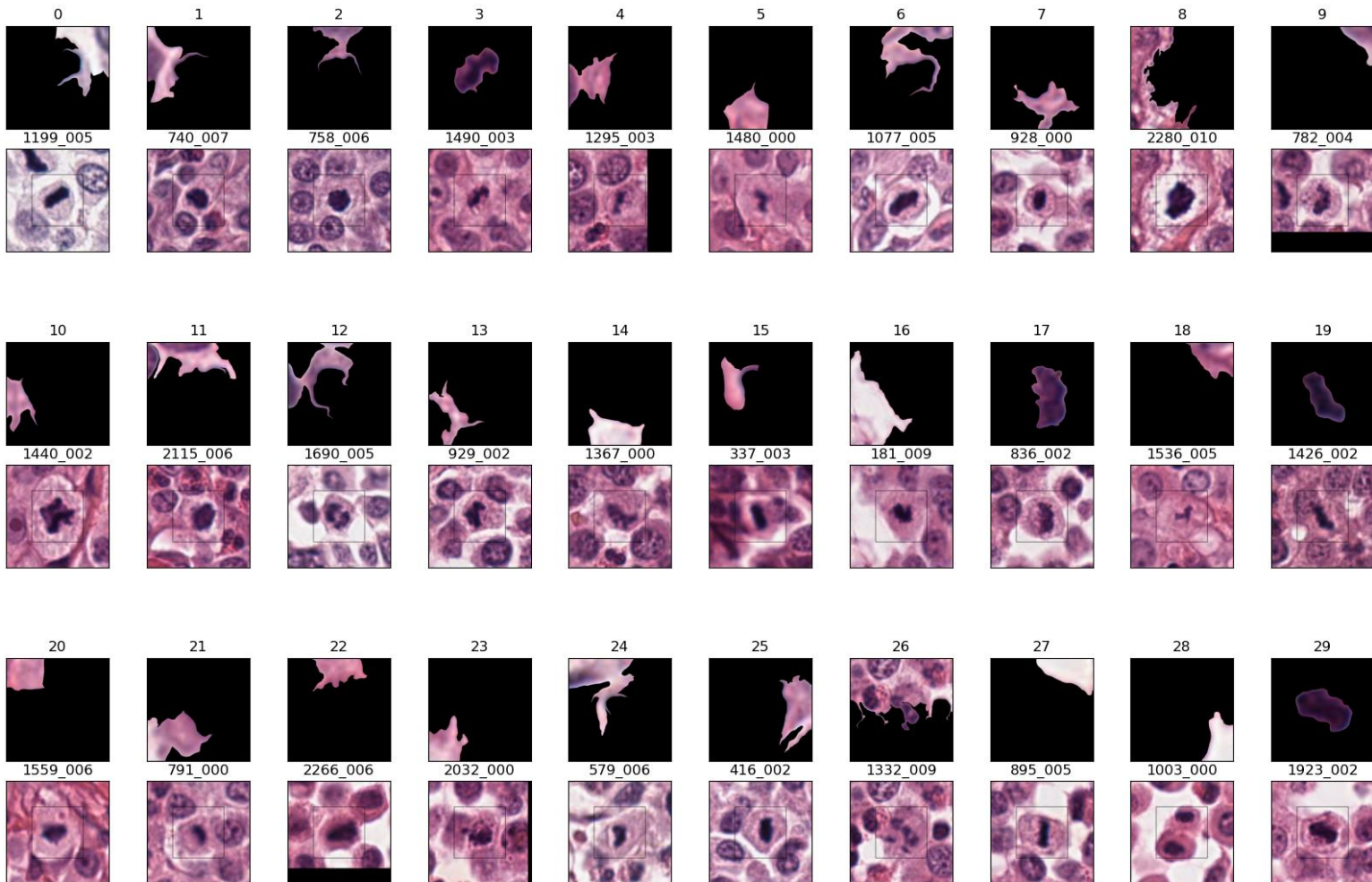backbone.body.layer4.2.conv1 mitotic figure_concept33

Q4b.

Are there any images in the above group that are not representative of this feature/concept? Please give the grid
index and briefly explain.

Canine Lymphoma

Q5a.

Is there any visual feature that these images (or most of) represent? If not, is there any pattern visible? Please answer below.

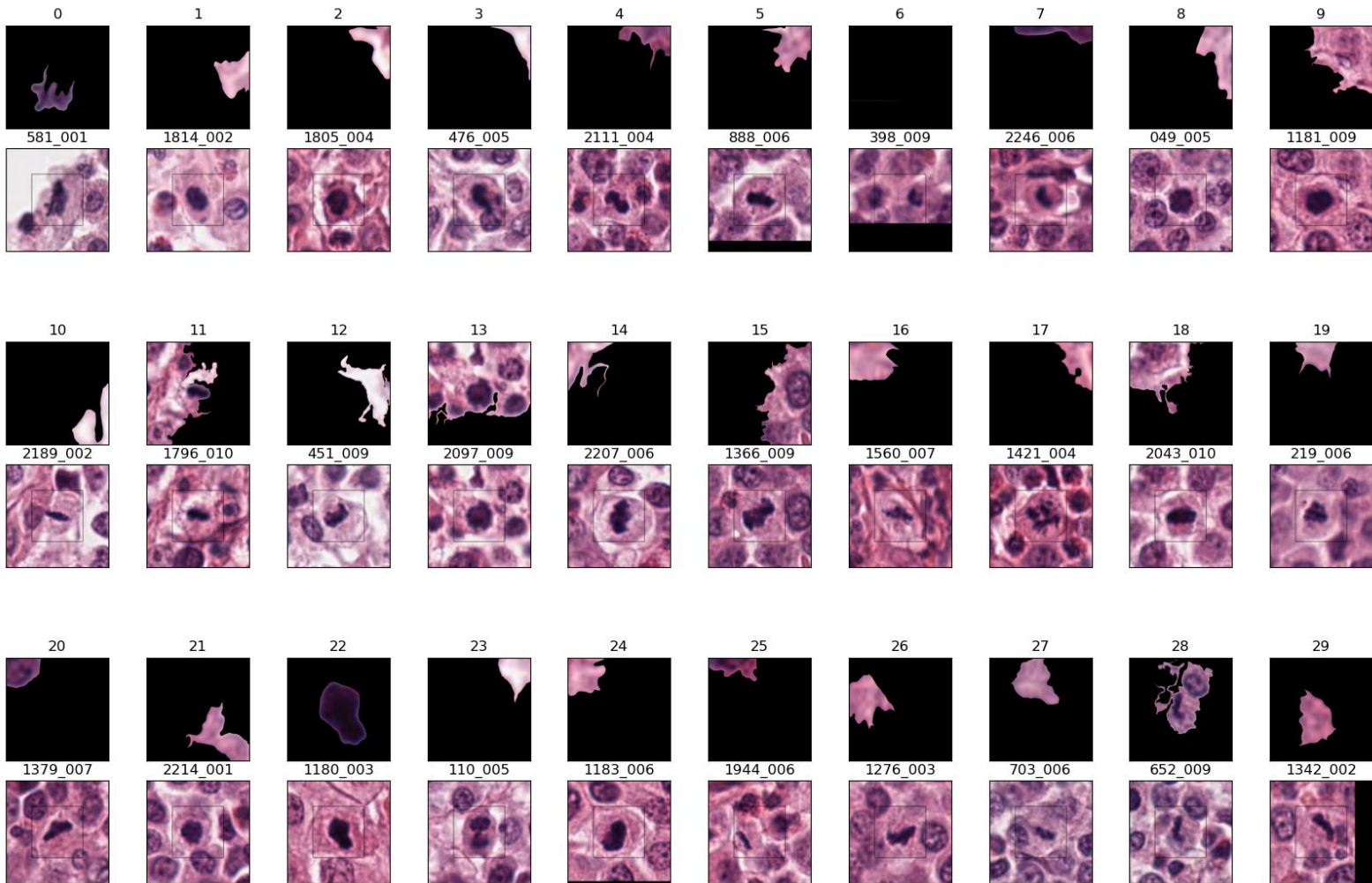backbone.body.layer4.2.conv1 mitotic figure_concept20

Q5b.

Are there any images in the above group that are not representative of this feature/concept? Please give the grid index and briefly explain.

Canine Lymphoma

Q6a.

Is there any visual feature that these images (or most of) represent? If not, is there any pattern visible? Please answer below.

backbone.body.layer4.2.conv1 mitotic figure_concept17

Q6b.

Are there any images in the above group that are not representative of this feature/concept? Please give the grid index and briefly explain.

Canine Cutaneous Mast Cell Tumour

Q7a.

Is there any visual feature that these images (or most of) represent? If not, is there any pattern visible?
Please answer below.

backbone.body.layer4.2.conv1 mitotic figure_concept32

Q8b.

Are there any images in the above group that are not representative of this feature/concept? Please give the grid index and briefly explain.
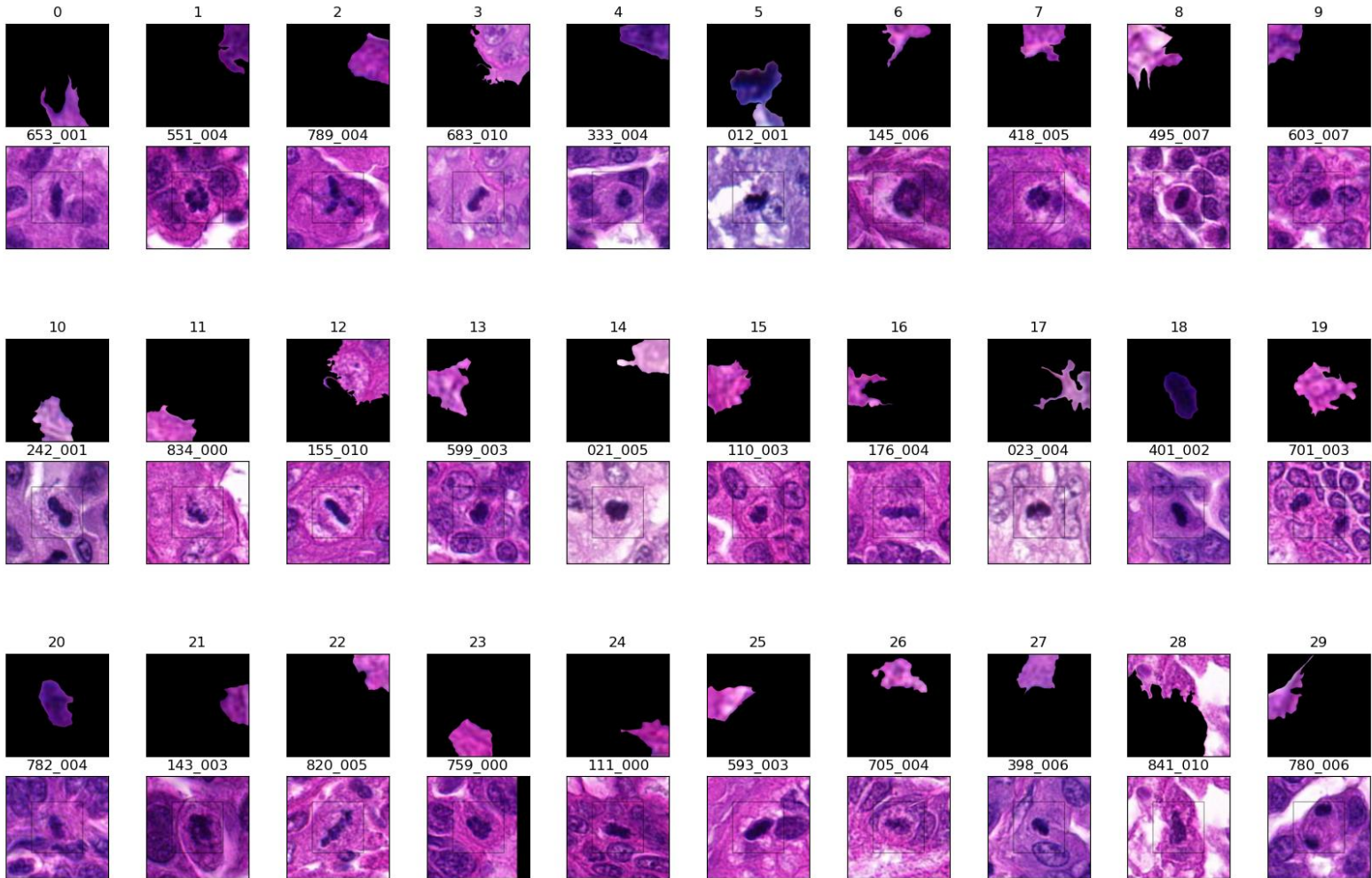
Canine Cutaneous Mast Cell Tumour

Q8a.

Is there any visual feature that these images (or most of) represent? If not, is there any pattern visible?
Please answer below.

backbone.body.layer4.2.conv1 mitotic figure_concept49

Q8b.

Are there any images in the above group that are not representative of this feature/concept? Please give the grid index and briefly explain.

Q9a.

Is there any visual feature that these images (or most of) represent? If not, is there any pattern visible?
Please answer below.

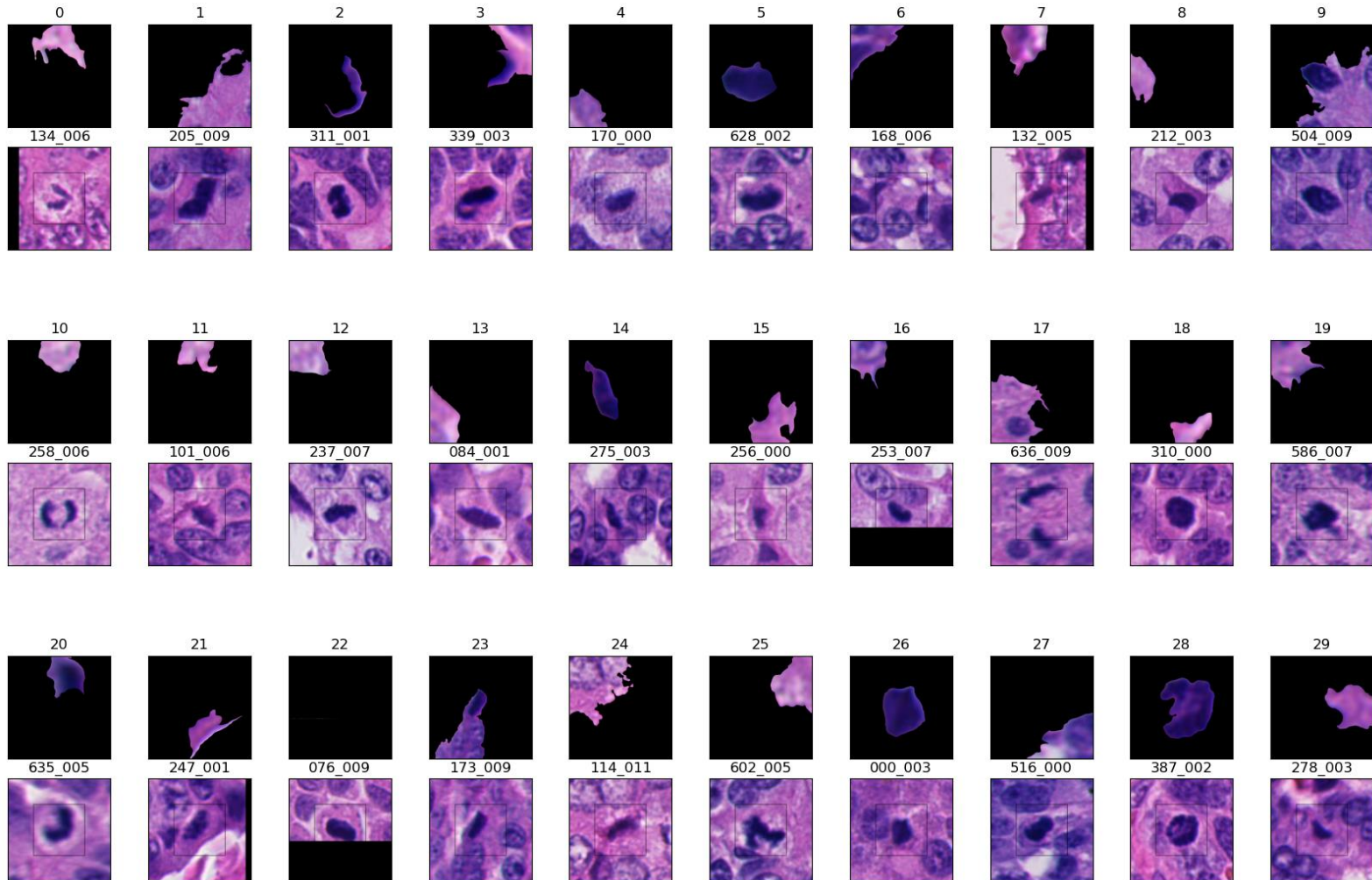backbone.body.layer4.2.conv1 mitotic figure_concept15

Q9b.

Are there any images in the above group that are not representative of this feature/concept? Please give the grid
index and briefly explain.

Human Neuroendocrine Tumour

Q10a.

Is there any visual feature that these images (or most of) represent? If not, is there any pattern visible?
Please answer below.

backbone.body.layer4.2.conv1 mitotic figure_concept46

Q10b.

Are there any images in the above group that are not representative of this feature/concept? Please give the grid index and briefly explain.
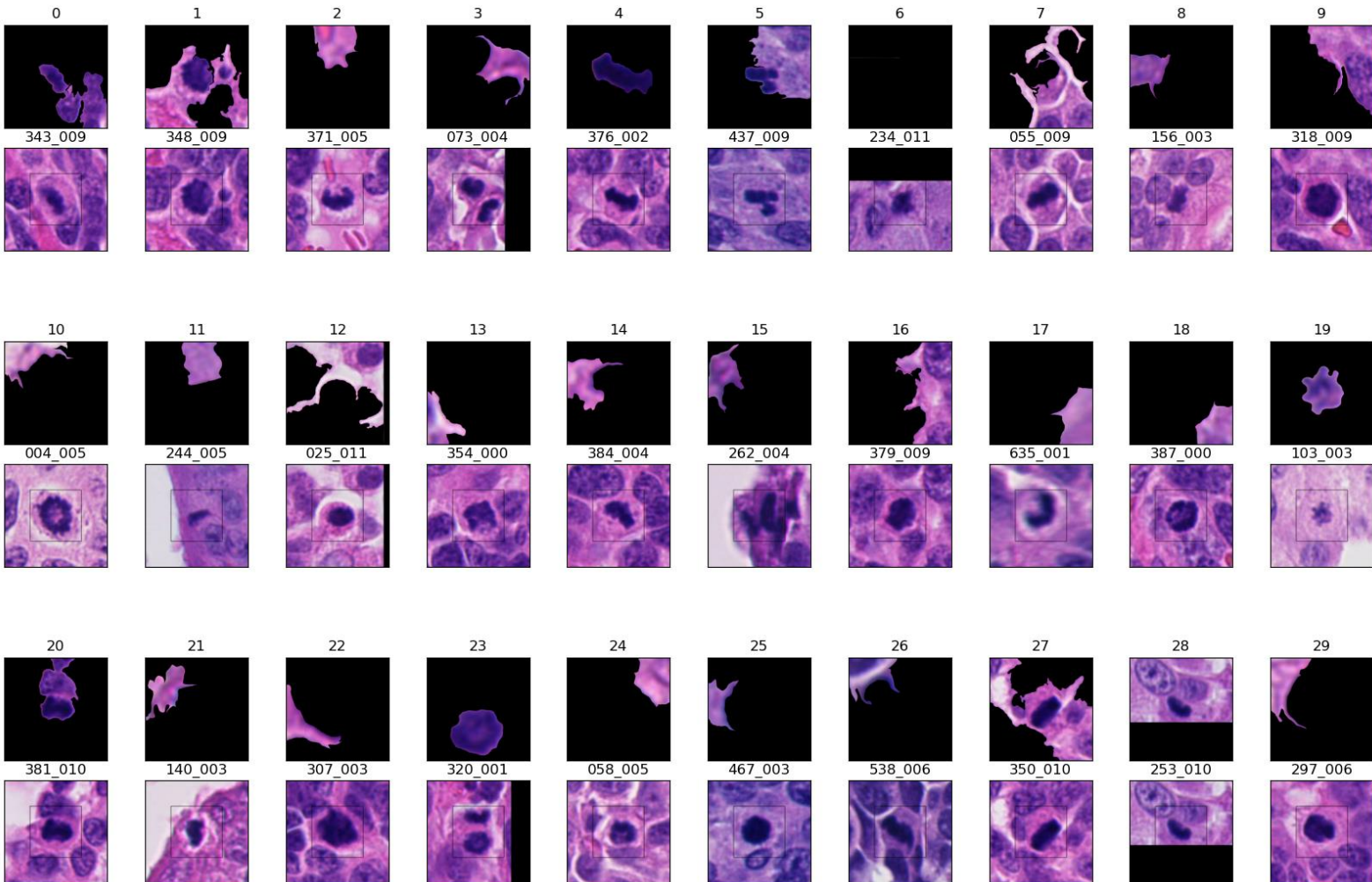
Human Neuroendocrine Tumour

Q11a.

Is there any visual feature that these images (or most of) represent? If not, is there any pattern visible?
Please answer below.

backbone.body.layer4.2.conv1 mitotic figure_concept22

Q11b.

Are there any images in the above group that are not representative of this feature/concept? Please give the grid index and briefly explain.

This is the end of questionnaire. Thank you for taking the time to review these results!