

# GitHubにおけるバグ報告等の 動画及び画像の活用実態に関する調査

蔵元 宏樹<sup>†</sup> 石本 優太<sup>†</sup> 新堂 風<sup>†</sup>  
近藤 将成<sup>†</sup> 柏 祐太郎<sup>†</sup> 亀井 靖高<sup>†</sup>  
鵜林 尚靖<sup>†</sup>

<sup>†</sup> 九州大学

E-mail: <sup>†</sup>{kuramoto,ishimoto,shindo}@posl.ait.kyushu-u.ac.jp,

<sup>††</sup>{kondo,kashiwa,kamei,ubayashi}@ait.kyushu-u.ac.jp

あらまし バグの症状の迅速で的確な把握は、デバッグ作業の効率に大きく影響する。バグの再現動画やバグの状況を表す画像は、デバッグ作業に有用であると期待できる。本研究の目的は、ソフトウェア開発現場におけるバグ報告動画及び画像の活用実態を明らかにすることである。GitHubで公開されている約4,000件のリポジトリを対象に、課題報告（Issue レポート）において動画及び画像の有無による影響を調査した。その結果として、動画及び画像が含まれる課題報告は、そうでないものに比べて、課題報告が解決するまでの時間が平均で、2.9%~17.6% 増加し、またコメント数が平均で 76.4%~143% 増加した。一方で、最初のコメントがつくまでの時間が平均で 20.1%~25.4% 減少した。これらの結果の詳細について報告する。

キーワード GitHub, 動画, 画像, バグ報告, Issue

## The investigation of the usage of videos and images for bug reports in GitHub

Hiroki KURAMOTO<sup>†</sup>, Yuta ISHIMOTO<sup>†</sup>, Kaze SHINDO<sup>†</sup>,

Masanari KONDO<sup>†</sup>, Yutaro KASHIWA<sup>†</sup>, Yasutaka KAMEI<sup>†</sup>, and Naoyasu UBAYASHI<sup>†</sup>

<sup>†</sup> Kyushu University

E-mail: <sup>†</sup>{kuramoto,ishimoto,shindo}@posl.ait.kyushu-u.ac.jp,

<sup>††</sup>{kondo,kashiwa,kamei,ubayashi}@ait.kyushu-u.ac.jp

**Abstract** A quick and accurate understanding of the symptoms of a bug can greatly affect the efficiency of debugging. Videos of bug reproductions and images showing the status of bugs are expected to be useful in the debugging process. The purpose of this study is to clarify the actual usage of bug report videos and images in software development sites. In this study, we investigated the impact of including video or images in the issue reports of about 4,000 public repositories in GitHub. As a result, the average time to resolve an issue report increased by 2.9%~17.6% and the average number of comments increased by 76.4%~143% for issue reports that included video and images compared to those that did not. On the other hand, the average time until the first comment decreased by 20.1%~25.4%. We report the details of these results.

**Key words** GitHub, Movie, Image, Bug report, Issue

### 1. はじめに

ソフトウェア開発を行う際の重要な工程の一つとしてデバッグ作業があげられる。デバッグ作業がソフトウェ

アの開発コストの 50% 以上を占めるという結果が示されている [1], [2]。この問題を解決するためにバグ修正効率化の研究が盛んに行われている。

バグの症状の迅速で的確な把握は、バグの再現性を高

め、デバッグ作業の効率に大きく影響する点で重要であると考えられる。バグの再現動画及びバグの状況を表す画像は、バグの症状把握に有用であり、これを有効に活用することでデバッグ作業の効率化が期待できる。

本研究の目的は、ソフトウェア開発現場におけるバグ報告動画及び画像の活用実態を明らかにすることである。GitHubで公開されている約4,000件のリポジトリを対象に、課題報告（Issue レポート）において動画及び画像の有無による影響を調査した。その調査結果について、報告する。

本稿では、2節で本研究を行うに至った動機及び調査課題を述べ、3節で本調査で用いる分析法について述べ、4節で本研究で用いるデータセットについて述べる。5節で調査内容及び調査結果を述べ、6節で調査課題の回答と考察を述べる。7節で妥当性への脅威について述べ、最後に8節でまとめと今後の課題について述べる。

## 2. 動 機

GitHubでは、公開されたリポジトリに対して、そのオーナーだけでなく不特定多数のユーザーが開発に貢献しており、不具合や意見の報告時に Issue 機能が使用されている。2021年以前は、Issue 作成には文字入力他に、GIF 動画や画像のみ添付可能であり、動画添付時には一度 GIF 動画に変換する手間があった。現在では、MP4、MOV ファイルを容易に添付できるようになり、今後動画の活用件数は増加すると考えられる。しかし現時点では、動画及び画像の有無による Issue への影響について明らかにされていない。本調査では、次の3つの RQ を設定し、回答する。

**RQ1：動画及び画像の活用頻度はどの程度か。**

RQ1では、動画及び画像のどの程度の頻度で用いられているかを明らかにする。

**RQ2：動画及び画像の期待される効果は何か。**

RQ2では、動画及び画像の有無が、解決までの時間やコメントの回数など Issue のパラメータにどのように影響するかを明らかにする。

**RQ3：どのような報告内容に動画及び画像を用いるか。**

RQ3では、動画及び画像がどのような単語と併用されているかを調査することで、どのような報告内容のときに動画及び画像が用いられているかを明らかにする。

## 3. 本調査で用いる分析法

本節では、本調査で用いた検定法及び指標について紹介する。

### 3.1 検 定 法

本調査では、画像を含む Issue、動画を含む Issue、ど

ちらも含まない Issue の3群の比較を行う。差の検定法では一般に、 $t$  検定が用いられるが、3群のそれぞれの組み合わせに対して1回あたり有意水準  $\alpha$  で検定を3回繰り返すと、真の有意水準が  $\alpha$  よりも大きくなってしまふ。誤って有意差があると判断してしまうことにつながる。3群以上の群間比較では、比較回数に応じて有意水準または有意確率を調節する多重比較が行われる。本調査では、正規性及び等分散性を満たさないデータにも対応する Steel-Dwass 法を採用する。Steel-Dwass 法は、Wilcoxon の順位和検定に基づくノンパラメトリックな多重比較法である。Wilcoxon の順位和検定は、以下の概念による検定法である。データがいずれの群のものであるかに関わらず、全てのデータに対して順位を付ける。各群ごとに順位を合計し、その順位和を比較する。帰無仮説は「すべての群の母集団分布は同じ」であり、棄却された時の解釈は「同じ母集団から採られたものではない」となる。ここで注意しなければならない点は、「群 A は群 B よりも有意に大きい」と解釈できるのは、すべての群を通して等分散性が成立するときである。[3] また、順位和検定は順位を用いるため、平均値よりも中央値付近に強く影響を受ける。その他、正規性の検定法には、Kolmogorov-Smirnov 検定を採用し、等分散性の検定法には、Levene 検定を採用する。

標本化について。比較する各群から同一サイズの標本を抽出し、有意確率を算出する試行を10回行う。10回の平均値を検定で得た有意確率とする。サンプルサイズは、用いる有意水準、検出力、効果量及び用いる検定法によって決まる。有意水準は一般的な0.05とする。その他の水準には、やや恣意的に検出力を0.80、効果量を0.10とする。これらと、3群の対応のない比較の条件のもとで、各群から323サンプルとする。サンプルサイズの算出には **G\*Power** [4] [5] を利用した。効果量は一般に先行研究からわかっている効果量を用いるが、本調査ではそれが得られなかった。従って、0.10（効果量小）、0.25（効果量中）、0.40（効果量大）の **Cohen (1988)** の基準から小程度の効果量を採用した。なお、多重比較法における検出力の考え方についての枠組みは、サンプルサイズの設計という観点では、明確化されていない様である。[6] 効果量、及び、検出力をやや恣意的に決定していることによるサンプルサイズへの影響は、6節で考察する。

### 3.2 tf\_idf

tf\_idf とは、文書中に含まれる単語の重要度を評価する指標である。tf とはある語彙の出現頻度であり、idf とはある語彙の出現する文書数の逆数を取ったものである。tf\_idf は tf と idf を掛け合わせたものである。語彙  $t$  が文書  $d$  に出現する回数を  $f$ 、 $d$  の全単語数（重複

を許す)を  $T$ ,  $t$  が出現する文書数を  $n$ , 全文書数を  $N$  とすると,  $d$  内の  $t$  に対する  $tf\_idf$  は以下の式で求められる.

$$\begin{aligned}tf(t,d) &= f/T \\idf(t) &= \log(N/n) \\tf\_idf(t,d) &= tf(t,d) * idf(t)\end{aligned}$$

$tf\_idf$  値が大きい単語は, その文書において重要な意味を持つと考えられる. Issue 群ごとに  $tf\_idf$  値が大きい単語を算出し, どのような報告内容に対して動画及び画像が用いられやすいかを分析する.

#### 4. データセット

本節では, 本研究で用いるデータセットについて紹介する.

##### 4.1 本調査の対象

次の条件を満たすリポジトリを対象とした.

- スターが 10 個以上あること
- 最終更新が 2021 年以降であること
- issue が一つ以上あること

スター数の条件は, 本研究に関連する可能性が低いプロジェクト (個人のプロジェクトやサンプルプロジェクトなど, 不具合報告が発生しにくいもの) を除外するために設定した. また, 最終更新を 2021 年以降に限定することで, より近年の動向を得られると判断した. すべての条件を満たすリポジトリは, 289,115 件であった. これらのリポジトリから 4,173 件のリポジトリ<sup>(注1)</sup>をランダムに選んだ. 4,173 件のリポジトリから, 770,656 件の解決済みの Issue を取得した. データ収集期間は 2021 年 11 月から 12 月であり, 件数は当時のものである.

##### 4.2 データ取得内容と方法

本調査で用いる Issue の 8 つの調査項目を紹介する. 調査項目を表 1 に示す.

Issue の調査項目は, GitHub API である PyGithub (python ライブラリ) を用いて取得した.  $Num\_of\_char$ ,  $Num\_of\_img$ ,  $Num\_of\_mov$ ,  $Words$  は直接取得できないため, その取得方法を紹介する. Issue に貼り付けた動画及び画像は URL に変換され, Issue のテキスト (直接取得可能) にマークダウン形式で記述される. URL に変換された一例を次に示す.

<https://user-images.githubusercontent.com/XXX/mp4>

XXX の部分は, 半角英数字, "/", "-" のみで構成される.

(注1): 本調査はデータ収集に十分な期間を確保できなかった為, 収集したリポジトリは全体の 2% 以下と少ない.

表 1 Issue の調査項目

| 調査項目                         | 内容                   |
|------------------------------|----------------------|
| <i>Issue_open_time</i>       | Issue が解決するまでの時間 (日) |
| <i>First_comment_time</i>    | 最初のコメントがつくまでの時間 (日)  |
| <i>Num_of_comments</i>       | 寄せられたコメント数           |
| <i>Num_of_char</i>           | Issue の作成時の文字記述量     |
| <i>Num_of_img</i>            | Issue の作成時の画像添付数     |
| <i>Num_of_mov</i>            | Issue の作成時の動画添付数     |
| <i>Words</i>                 | Issue の作成時の記述された英単語  |
| <i>Issue_created_at_year</i> | Issue が作成された年        |

従って, 以上の URL を正規表現で取得し出現回数を, 拡張子に応じて  $Num\_of\_img$  及び  $Num\_of\_mov$  とした.

動画及び画像を含む Issue について, 上記 URL の除外処理を行ったテキストを用いて,  $Words$  を正規表現で取得した. また, 処理済みの Issue のテキストの長さを,  $Num\_of\_char$  とした.

我々は事前調査で,  $Issue\_open\_time$  が, 何らかの要因により負値をとる Issue が混在していることを発見した. このような不正な値を含む Issue を除外するため,  $30seconds \leq Issue\_open\_time \leq 1year$  の条件を設定した. この条件を満たす Issue は, 711,160 件 (92.23%) であった.

##### 4.3 データの分類

動画及び画像の有無で Issue を次の 3 つのカテゴリに分類した. 各 Issue 群の要素数を表 2 に示す. ただし, 動画と画像をどちらも含む Issue は  $Img$  と  $Mov$  の両方に属するものとする.

表 2 カテゴリ分類・要素数

|             | 説明                           | 要素数              |
|-------------|------------------------------|------------------|
| <i>Img</i>  | $Num\_of\_img$ が 1 以上の Issue | 33,079 (4.65%)   |
| <i>Mov</i>  | $Num\_of\_mov$ が 1 以上の Issue | 3,819 (0.54%)    |
| <i>None</i> | 上記どちらにも含まれない Issue           | 674,793 (94.81%) |

収集したデータの一部抜粋を表 3 に示す. 1 つの Issue の情報が表の行に並んでおり,  $Issue\_open\_time$  及び  $First\_comment\_time$  の単位は日数である.

## 5. 調 査

### 5.1 動画・画像の活用割合

図 1 は, 動画及び画像の活用割合の推移である. 2016 年までは, 画像の活用は 0.01% 未満であり, 動画に関しては 0 件であった. 2017 年以降, 活用件数は急激に増加し, 2021 年には, 動画 0.72%, 画像 5.93% であった. 2017 年に急激に増加した要因に関しては調査中である.

表 3 取得した Issue の調査項目の一部抜粋

| Issue_created<br>_at_year | Issue_open<br>_time | Num_of<br>_img | Num_of<br>_mov | Num_of<br>_comments | First_comment<br>_time | Num_of<br>_char |
|---------------------------|---------------------|----------------|----------------|---------------------|------------------------|-----------------|
| 2020                      | 6.99861111          | 0              | 0              | 1                   | 6.99861111             | 39488           |
| 2020                      | 41.9594329          | 1              | 0              | 3                   | 17.7784722             | 950             |
| 2020                      | 43.8850579          | 0              | 0              | 2                   | 0.49828704             | 379             |
| 2020                      | 44.0935532          | 0              | 0              | 4                   | 0.91277778             | 191             |
| 2020                      | 0.14934028          | 0              | 0              | 8                   | 0.08077546             | 1800            |
| 2020                      | 59.5670949          | 2              | 0              | 5                   | 0.39472222             | 999             |
| 2020                      | 74.9322569          | 0              | 0              | 0                   | -                      | 127             |

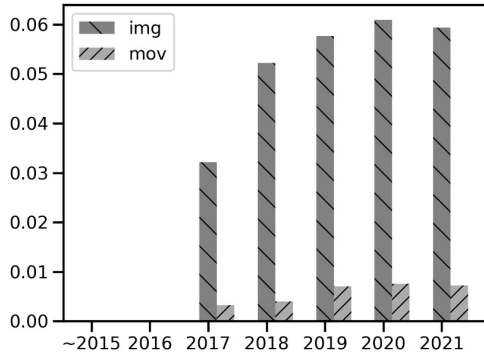


図 1 動画・画像の活用割合推移

## 5.2 データの分析

図2は、データの特徴値を箱ひげ図及び表で表現したものである。ただし、図2の箱図の最大値は、 $3rd\ Qu+1.5 \times IQR$  以下の最大値である。(IQR：四分位範囲)

動画及び画像を含む Issue は、そうでない Issue に比べて、*Issue\_open\_time* が平均で、2.9%~17.6% 増加し、*Num\_of\_comments* が平均で 76.4%~143% 増加した。一方で、*First\_comment\_time* が平均で 20.1%~25.4% 減少し、*Num\_of\_char* が平均で 26.3%~37.5% 減少した。正規性について、Kolmogorov-Smirnov 検定の結果、すべての項目に対して、正規性は有意水準 0.05 で棄却された。

等分散性について、Leneve 検定の結果を表4に示す。帰無仮説は「すべての Issue 群間を通して等分散」であり、有意水準は 0.20 とする。

表 4 等分散性検定の結果

|                           | 有意確率    | 平均分散比 |
|---------------------------|---------|-------|
| <i>Issue_open_time</i>    | 0.378   | 1.050 |
| <i>First_comment_time</i> | 0.296   | 1.301 |
| <i>Num_of_comments</i>    | * 0.001 | 2.459 |
| <i>Num_of_char</i>        | * 0.073 | 3.156 |

\* : 有意水準 0.20 で有意

*Num\_of\_comments* 及び *Num\_of\_char* は、等分散性が有意水準 0.20 で棄却される。*Issue\_open\_time* 及

表 5 調査項目ごとの多重比較の結果

| 調査項目                      | 対                  | 有意確率     |
|---------------------------|--------------------|----------|
| <i>Issue_open_time</i>    | <i>Img vs None</i> | ** 0.002 |
|                           | <i>Mov vs None</i> | ** 0.021 |
|                           | <i>Img vs Mov</i>  | 0.381    |
| <i>First_comment_time</i> | <i>Img vs None</i> | 0.764    |
|                           | <i>Mov vs None</i> | 0.351    |
|                           | <i>Img vs Mov</i>  | 0.404    |
| <i>Num_of_comments</i>    | <i>Img vs None</i> | * 0.001  |
|                           | <i>Mov vs None</i> | * 0.001  |
|                           | <i>Img vs Mov</i>  | 0.211    |
| <i>Num_of_char</i>        | <i>Img vs None</i> | * 0.001  |
|                           | <i>Mov vs None</i> | * 0.001  |
|                           | <i>Img vs Mov</i>  | 0.599    |

\* : 両側検定で有意 \*\* : 片側検定で有意

び *First\_comment\_time* は、有意確率が有意水準以上であることと、平均分散比が 1 ~ 1.5 程度であることから、等分散性を仮定する。

## 5.3 統計的検定

ここでの検定の目的は、各調査項目の群間の分布の差が統計的に有意であることを確かめることである。5.2 節において、正規性及び等分散性を満足しない調査項目が含まれることを確かめた。従って、検定には、正規性及び等分散性を仮定しない Steel-Dwass 法を用いる。Steel-Dwass 法による検定結果を表5に示す。有意水準は 0.05 とする。

*Img* 及び *Mov* は、*None* に比べて、*Issue\_open\_time* が有意に大きく、*Num\_of\_comments* 及び *Num\_of\_char* については分布に何らかの有意な差があると解釈できる。一方で、*First\_comment\_time* においては、有意差は認められなかった。また、*Img* と *Mov* の比較では、すべての項目で有意差は認められなかった。

## 5.4 出現単語分析

ここでの目的は、各 Issue 群の Issue にどのような単語が含まれているかを明らかにすることである。各 Issue の出現単語を *tf\_idf* 値に変換し、Issue 群ごとに結合した。

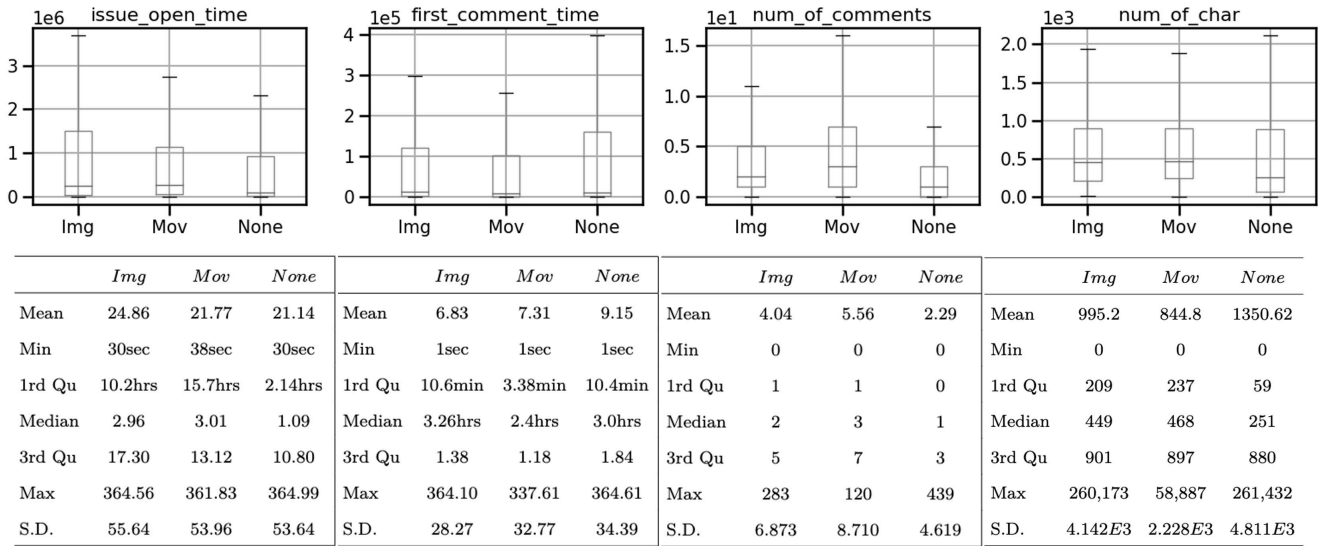


図 2 Issue 群ごとの調査項目の分布及び代表値

表 6 Issue 群ごとの **tf\_idf** 上位 10 単語

|    | Img        | Mov      | None        |
|----|------------|----------|-------------|
| 1  | image      | when     | dependabot  |
| 2  | screenshot | dropdown | code        |
| 3  | when       | view     | file        |
| 4  | error      | package  | pullrequest |
| 5  | screen     | issue    | version     |
| 6  | version    | python   | error       |
| 7  | shot       | height   | use         |
| 8  | file       | react    | lib         |
| 9  | test       | button   | add         |
| 10 | code       | text     | commit      |

各 Issue 群の **tf\_idf** 値上位 10 単語を表 6 に示す。"at" や "it", "the" などの単語を除外する為、名詞、動詞及び疑問詞のみを対象とした。

*Img* では "image" や "screen" など表示に関係する単語が多く、*Mov* では "dropdown" や "button" など動きのある表示に関する単語が多い。一方で、*None* では、リポジトリの管理に関する単語が多く、プルリクエストやコミットに関する通知報告がなされていると推察される。*Mov* では "when" が 1 番目に、*Img* でも 3 番目にランクインした。これは、不具合の再現に関する報告である可能性が高い。実際に、目視調査で、不具合の再現動画が多いことを実験的に確かめた。一方で、*None* では 1 番目に "dependabot" がランクインした。Dependabot は、最新のライブラリを推奨したり更新するボットであり、設定により自動で Issue を生成することもある。*Img* 及び *Mov* と *None* の差は、ボットによる影響を受けている可能性がある。

## 6. RQ の回答と考察

本節では、5 節の調査結果から RQ1 RQ3 について回答

し、サンプルサイズについて考察する。

5.1 節の結果から、動画及び画像の活用頻度として以下の結果を得た。

### RQ1：動画及び画像の活用頻度はどの程度か。

- ・2016 年以前はほとんど用いられてこなかった。
- ・2017 年に急激に増加し、以降増加傾向である。
- ・2021 年には、動画 0.72%、画像 5.93% であった。

5.2 節では、動画及び画像の有無で、Issue のパラメータがどのように変化するかを分析した。また、5.3 節では、各調査項目の群間の分布の差が統計的に有意であることを確かめた。動画及び画像を含む Issue は、そうでない Issue に比べて、*Num\_of\_char* が平均で 26.3%~37.5% 減少したことから、動画及び画像を用いることでバグの症状を細かく記述せずに済むと推察される。また、*First\_comment\_time* が平均で 20.1%~25.4% 減少し、*Num\_of\_comments* が平均で 76.4%~143% 増加したことから、動画及び画像のある Issue は開発者の着手を早め、コメントを返信が得やすくなると推察される。また、*Img* と *Mov* の比較検定では、すべての項目で有意差が認められなかったことから、動画と画像の効果に有意な差はないと推察される。

### RQ2：動画及び画像の期待される効果は何か。

- ・バグの症状を細かく記述せずに済む。
- ・開発者の着手を早め、コメントを得やすい。
- ・動画と画像の効果に有意な差はない。

5.4 節では、出現単語分析を行うことで、どのような報告内容の場合に動画及び画像を用いるかを確かめた。*Img* では表示に関係する単語が多く、*Mov* では動きのある表

示に関する単語が多く見られた。一方で *None* では、プルリクエストやコミットなどリポジトリ管理に関する単語が多かった。

### RQ3：どの様な報告に動画及び画像を用いるか。

- ・表示に関する問題に対して、画像を用いる。
- ・動きのある表示に関する問題に対して、動画を用いる。

一般にサンプルサイズが多ければ多いほど、母集団の特徴をより正確に抽出できる。ところが検定においては、サンプルサイズが大きくなるにつれて、誤差に過敏になり有意確率が小さくなる傾向がある。すなわち、検定で有意差を得たとしても、母集団分布の有意な差によるものか、サンプルサイズが大きすぎることに由来のものか判断できない。従って、適切なサンプルサイズを前もって決めておくことが重要である。本調査では、事前に効果量の目安が得られなかったために、やや恣意的に決定した。検定結果を見ると、有意差がはっきりと認められた部分と、認められなかった部分が見られた *Issue\_open\_time*, *Num\_of\_comments* 及び *Num\_of\_char* に対しては適切なサンプルサイズであったと推察する。一方で *First\_comment\_time* に対しては、実際に有意と言える差が無い、あるいは、サンプルサイズが小さすぎた可能性がある。

## 7. 妥当性への脅威

内的妥当性。本調査で用いた等分散性の検定の有意水準について、留意しなければならない点がある。多重比較法は等分散性に関して、二群比較よりも鋭敏であり、有意水準 0.20 程度で行わなければ意味をなさないことが示されている。[7] さらには、有意水準 0.50 まで引き上げても、誤った結果を招く場合があると言われている。その場合、*Issue\_open\_time* 及び *First\_comment\_time* の等分散性は棄却され、「動画及び画像がある Issue は、そうでないものに比べて *Issue\_open\_time* が有意に大きい」とした表現は、「動画及び画像がある Issue は、そうでないものに比べて *Issue\_open\_time* に何らかの有意な差をもたらす」となる。

また、ボットが *None* に多くあらわれていることが、出現単語分析で明らかになった。bot を取り除く処理をしなければ、分布差が動画及び画像によるものとは断定できない。その点が考慮されていない。

また、本調査はデータ収集に十分な期間を確保できなかった為、収集したリポジトリは全体の 2% 以下と少ない。より正確な結果を得るために、収集量を増やすことは今後の課題である。

外的妥当性。本調査の調査対象は、条件に適合したリポジトリのみであり、すべての開発現場で同様の傾向があるとは限らない。また、開発規模や分野など、リポジトリの性質が考慮されていない。

## 8. まとめと今後の課題

本研究では、ソフトウェア開発現場におけるバグ報告動画及び画像の活用実態を明らかにするため、GitHub で公開されているいくつかのリポジトリを対象に、Issue と動画及び画像の有無の関係を調査した。

スターが 10 個以上のリポジトリに対して Issue 報告を行うのは、一般に開発中級者以上であると考えられる。本調査は中級者以上のバグ報告の実態調査であり、この試みは開発初級者がバグ報告する際の参考になると考えられる。

また、今後の課題として、次の 3 つの追加調査を計画している。

- ・ボットを可能な限り除外すること
- ・開発規模の大小で分類することで、実態をより細かく調査すること
- ・対象のリポジトリを増やし、より全体の傾向を結果に反映すること

## 謝 辞

本研究の一部は XXX の助成を受けた。

## 文 献

- [1] James S. Collofello and Scott N. Woodfield. Evaluating the effectiveness of reliability-assurance techniques. *Journal of Systems and Software*, Vol. 9, No. 3, pp. 191–195, 1989.
- [2] L. Gazzola, D. Micucci, and L. Mariani. Automatic software repair: A survey. *IEEE Transactions on Software Engineering*, Vol. 45, No. 1, pp. 34–67, 2019.
- [3] 山本光司. 植物防疫基礎講座 正しい分散分析結果を導くための変数変換法. 第 56 巻, pp. 436–441, 2002-10.
- [4] Lang. Albert-Georg Faul. Franz, Erdfelder. Edgar and Buchner. Axel. G\*power 3: A flexible statistical power analysis program for the social; behavioral; and biomedical sciences. Vol. 39, pp. 175–191, 2007.
- [5] Buchner. Axel Faul. Franz, Erdfelder. Edgar and Lang. Albert-Georg. Statistical power analyses using g\*power 3.1: Tests for correlation and regression analyses. Vol. 41, pp. 1149–1160, 2009.
- [6] 永田靖. 多重比較法の実践. 第 27 巻, pp. 93–108, 1998.
- [7] 阿部研自. 多重比較法における不等分散の影響評価. 第 28 巻, pp. 55–78, 1999.